

# UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS  
Departamento de Bioquímica y Biología Molecular I



## TESIS DOCTORAL

**Caracterización de proteínas de síntesis, degradación y unión de  
peptidoglicano: implicaciones en mecanismos de virulencia y  
resistencia a antibióticos en *S. pneumoniae***

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Javier Gutiérrez Fernández**

Directores

Juan Antonio Hermoso Domínguez

**Madrid, 2015**



---

**Caracterización de proteínas de síntesis,  
degradación y unión de peptidoglicano.  
Implicaciones en mecanismos de virulencia y  
resistencia a antibióticos en *S. pneumoniae*.**

---

Javier Gutiérrez Fernández

---

***Universidad Complutense de Madrid***  
***Consejo Superior de Investigaciones Científicas***

---



UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS BIOLÓGICAS

DEPARTAMENTO DE BIOQUÍMICA Y BIOLOGÍA MOLECULAR I



TESIS DOCTORAL

Caracterización de proteínas de síntesis, degradación y unión de peptidoglicano. Implicaciones en mecanismos de virulencia y resistencia a antibióticos en *S. pneumoniae*.

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Javier Gutiérrez Fernández**

Director

Juan Antonio Hermoso Domínguez

Madrid, 2015



El Dr. **Juan Antonio Hermoso Domínguez**, Profesor de Investigación del Departamento de Cristalografía y Biología Estructural del Instituto de Química-Física “Rocasolano” (Consejo Superior de Investigaciones Científicas, CSIC), certifica que:

Don Javier Gutiérrez Fernández, Licenciado en Biotecnología por la Universidad de León, ha realizado bajo mi dirección el trabajo titulado “Caracterización de proteínas de síntesis, degradación y unión de peptidoglicano. Implicaciones en mecanismos de virulencia y resistencia a antibióticos en *S. pneumoniae*”, en el Departamento de Cristalografía y Biología Estructural del Instituto de Química-Física “Rocasolano” (Consejo Superior de Investigaciones Científicas, CSIC).

En mi opinión, este trabajo reúne las condiciones exigidas por la legislación vigente y tiene la originalidad, el rigor y la calidad científica necesarios y suficientes para ser presentado por Don Javier Gutiérrez Fernández para optar al Grado de Doctor por la Universidad Complutense de Madrid.

Para que así conste y con los efectos oportunos, firmo el presente certificado.

Madrid, 23 de Enero 2015.

Fdo: Dr. Juan Antonio Hermoso Domínguez  
Profesor de Investigación, CSIC



Quisiera expresar mi más sincero agradecimiento al Dr. Juan A. Hermoso, director de esta tesis, por su apoyo y motivación durante estos cuatro años. Su forma de trabajar ha sido contagiosa y me ha animado a ir siempre más allá, abordando los problemas desde puntos de vista novedosos para mi, algo sin duda enriquecedor.

También quiero agradecer a esta tesis a la gente que ha estado o ha pasado por el Departamento de Cristalografía y Biología Estructural del Instituto de Química-Física “Rocasolano” durante mi estancia, ya que me he beneficiado de la interacción con todos y cada uno de ellos. Gran parte de la formación con la que termino esta etapa se la debo a las numerosas charlas y discusiones que hemos entablado a lo largo de estos años.

Quiero además hacer una mención especial a la gente de fuera del Instituto con la que he tenido una interacción más que productiva. Las colaboraciones directas en las que he tenido la suerte de participar, especialmente con los grupos del Dr. Sven Hammerschmidt (Universidad de Greifswald, Alemania), de la Dra. María Jesús Martínez (Centro de Investigaciones Biológicas, CIB-CSIC) y de la Dra. Kathrin Mühlemann (Universidad de Berna, Suiza), me han aportado una visión global de los proyectos que hemos abordado, permitiéndome aprender nuevas técnicas y aproximaciones a un problema común. Además, las interacciones con el grupo de la Dra. Isabel Usón y con el Dr. Alfonso Martínez me han servido para afrontar ciertos aspectos específicos de la técnica cristalográfica.

Por supuesto no puedo olvidarme del equipo de Xaloc y la gente de ALBA (especialmente los Dres. Jordi Juanhuix, Fernando Gil y Daniel Fulla), tanto por su ayuda a la hora de realizar los experimentos de difracción como por su motivación y apoyo personal en las numerosas visitas que he tenido la suerte de realizar.

Finalmente, quienes más se merecen este agradecimiento son amigos y familia, sin cuyo apoyo no hubiese podido concluir esta etapa. Los primeros, por mantenerse ahí durante todo este tiempo, haciendo esta andanza lo más agradable posible. Los segundos, por su constante ánimo, consejo y, sobre todo, paciencia. Muchas gracias a todos.





*A mis padres  
y a Marta*



---

# ÍNDICE GENERAL

---

ABSTRACT . . . . .	XV
<b>I INTRODUCCIÓN</b>	<b>1</b>
<b>1. <i>Streptococcus pneumoniae</i></b>	<b>3</b>
1.1 Características generales . . . . .	4
1.2 Estructura de la pared celular . . . . .	6
1.3 Remodelado de la pared . . . . .	10
1.3.1 Biosíntesis . . . . .	10
1.3.2 División celular . . . . .	13
1.3.3 Degradación de peptidoglicano . . . . .	14
1.4 Proteínas de superficie . . . . .	14
1.4.1 Proteínas con motivo LPXTG . . . . .	15
1.4.2 Lipoproteínas . . . . .	16
1.4.3 Proteínas no-clásicas . . . . .	16
1.4.4 Proteínas de unión a colina . . . . .	16
1.5 Utilización de antibióticos y mecanismos de resistencia . . . . .	18
1.5.1 Fenómeno de heterorresistencia . . . . .	19
<b>2. Objetivos</b>	<b>21</b>
<b>II MATERIAL Y MÉTODOS</b>	<b>23</b>
<b>3. Cristalografía de rayos X</b>	<b>25</b>
3.1 Obtención de la proteína . . . . .	26
3.1.1 Clonación y expresión . . . . .	27
3.1.1.1 Cepas bacterianas . . . . .	27
3.1.1.2 Vectores de expresión . . . . .	28
3.1.1.3 Medios de cultivo . . . . .	28
3.1.1.4 Expresión de proteínas . . . . .	29
3.1.2 Purificación . . . . .	29
3.1.2.1 Lisis celular . . . . .	29
3.1.2.2 Cromatografía de afinidad . . . . .	30

---

3.1.2.3	Cromatografía de intercambio iónico . . . . .	30
3.1.2.4	Cromatografía de exclusión molecular . . . . .	31
3.2	Cristalización de proteínas . . . . .	31
3.2.1	Nucleación y crecimiento cristalino . . . . .	31
3.2.2	Control del crecimiento cristalino . . . . .	32
3.2.3	Cristalización en presencia de ligandos . . . . .	33
3.2.4	Configuración de los experimentos de cristalización . . . . .	34
3.2.4.1	Difusión de vapor con aislamiento . . . . .	34
3.2.4.2	Difusión de vapor sin aislamiento . . . . .	35
3.2.4.3	Contradifusión en gel . . . . .	36
3.2.5	Búsqueda de condiciones iniciales y optimización . . . . .	36
3.2.6	Defectos en los cristales . . . . .	38
3.2.6.1	Cristales maclados . . . . .	38
3.3	Difracción de rayos X y resolución estructural . . . . .	39
3.3.1	Configuración del experimento de difracción . . . . .	39
3.3.1.1	Fuentes de rayos X . . . . .	39
3.3.1.2	Soporte y protección del cristal . . . . .	41
3.3.1.3	Detección de rayos X . . . . .	43
3.3.2	Obtención y análisis del patrón de difracción . . . . .	44
3.3.2.1	Dispersión de la radiación y factor de dispersión atómico . . . . .	44
3.3.2.2	Simetría, red directa y red recíproca . . . . .	45
3.3.2.3	Interacción geométrica de la luz con el cristal . . . . .	46
3.3.2.4	Esfera de Ewald y registro del patrón de difracción . . . . .	47
3.3.2.5	Información aportada por el patrón de difracción . . . . .	48
3.3.3	Determinación de la estructura tridimensional de la proteína . . . . .	49
3.3.3.1	Reducción de datos y cálculo de $ F_{hkl} $ . . . . .	50
3.3.3.2	El problema de las fases y cálculo de $\varphi_{hkl}$ . . . . .	53
3.3.3.3	Método de reemplazo molecular . . . . .	54
3.3.3.4	Faseado experimental de subestructuras . . . . .	55
3.3.3.5	Cálculo de la densidad electrónica, construcción del modelo y refinamiento . . . . .	60
3.3.3.6	Validación . . . . .	65
3.3.4	Detección, análisis y resolución de datos maclados . . . . .	65

### III RESULTADOS Y DISCUSIÓN 71

4.	Determinación estructural y funcional de MurA1 . . . . .	73
4.1	Caracterización de MurA1 . . . . .	74
4.2	Obtención de MurA1 <sub>D39</sub> y MurA1 <sub>Hungary</sub> <sup>19A</sup> . . . . .	74
4.3	Cristalización de MurA1 <sub>D39</sub> . . . . .	75
4.4	Determinación estructural de MurA1 <sub>D39</sub> . . . . .	76
4.5	Análisis estructural de MurA1 <sub>D39</sub> . . . . .	78
4.5.1	Dominios estructurales . . . . .	78

---

---

4.5.2	Empaquetamiento cristalino . . . . .	79
4.5.3	Agitación térmica . . . . .	80
4.5.4	Solvente . . . . .	80
4.6	Cristalización de MurA1 <sub>Hungary</sub> <sup>19A</sup> . . . . .	80
4.7	Determinación estructural de MurA1 <sub>Hungary</sub> <sup>19A</sup> . . . . .	81
4.8	Análisis estructural de MurA1 <sub>Hungary</sub> <sup>19A</sup> . . . . .	83
4.8.1	Estructura general . . . . .	83
4.8.2	Empaquetamiento cristalino . . . . .	84
4.8.3	Agitación térmica . . . . .	84
4.8.4	Otras moléculas y solvente . . . . .	85
4.9	Cristalización de MurA1 <sub>D39</sub> y MurA1 <sub>Hungary</sub> <sup>19A</sup> en presencia de ligandos . . . . .	86
4.10	Relación estructura-función . . . . .	87
4.10.1	Análisis comparado de estructuras . . . . .	87
4.10.2	Reconocimiento de sustratos y sitio activo . . . . .	89
4.10.3	Inhibición mediante fosfomicina y citrato . . . . .	91
4.10.4	Hipótesis sobre el mecanismo de reacción . . . . .	94
4.10.5	MurA1 <sub>Hungary</sub> <sup>19A</sup> y heterorresistencia a fosfomicina . . . . .	95
<b>5.</b>	<b>Determinación estructural y funcional de DacB</b>	<b>99</b>
5.1	Caracterización de DacB . . . . .	100
5.2	Cristalización y resolución parcial de DacB <sub>20–238</sub> . . . . .	100
5.2.1	Expresión y purificación . . . . .	100
5.2.2	Cristalización y optimización . . . . .	101
5.2.3	Determinación estructural . . . . .	102
5.2.4	Análisis estructural . . . . .	105
5.2.4.1	Estructura general y sitio de unión a zinc . . . . .	106
5.2.4.2	Empaquetamiento cristalino . . . . .	106
5.2.4.3	Agitación térmica . . . . .	108
5.2.4.4	Otras moléculas y solvente . . . . .	108
5.3	Segunda construcción (DacB <sub>61–238</sub> ) . . . . .	109
5.3.1	Expresión y purificación . . . . .	109
5.3.2	Cristalización y resultados preliminares . . . . .	109
5.3.3	Variantes de la cristalización para evitar maclas . . . . .	113
5.3.4	Análisis de maclas y determinación estructural . . . . .	116
5.3.5	Faseado experimental de cristales maclados . . . . .	122
5.3.6	Determinación de la existencia del sitio de unión a zinc . . . . .	124
5.3.7	Análisis estructural . . . . .	124
5.3.7.1	Estructura general . . . . .	125
5.3.7.2	Empaquetamiento cristalino . . . . .	126
5.3.7.3	Agitación térmica . . . . .	126
5.3.7.4	Otras moléculas y solvente . . . . .	127
5.4	Tercera construcción (DacB <sub>49–238</sub> ) . . . . .	127
5.4.1	Expresión y purificación . . . . .	127

---

---

5.4.2	Cristalización . . . . .	128
5.4.3	Determinación estructural . . . . .	128
5.4.4	Análisis estructural . . . . .	130
5.4.4.1	Estructura general . . . . .	130
5.4.4.2	Empaquetamiento cristalino . . . . .	131
5.4.4.3	Agitación térmica . . . . .	134
5.4.4.4	Átomos de zinc y solvente . . . . .	135
5.4.5	Relación entre DacB <sub>49–238</sub> y DacB <sub>61–238</sub> . . . . .	135
5.4.6	Cristalización en presencia de ligandos . . . . .	137
5.5	Relación estructura-función . . . . .	138
5.5.1	Maduración de DacB . . . . .	138
5.5.2	Análisis estructural comparado . . . . .	139
5.5.3	Maquinaria catalítica y conformaciones estructurales . . . . .	141
5.5.4	Estudios microbiológicos y fisiológicos relacionados con DacA y DacB . . . . .	145
5.5.4.1	Localización genómica de DacB . . . . .	146
5.5.4.2	Anclaje de DacA y DacB a la membrana y ensayos de expresión . . . . .	146
5.5.4.3	Caracterización fenotípica de células deficientes en DacA y DacB . . . . .	146
5.5.4.4	Influencia en la composición del peptidoglicano . . . . .	147
5.5.4.5	Modelos de infección en ratones y supervivencia en macrófagos . . . . .	149
5.5.4.6	Adherencia . . . . .	149

## 6. Determinación estructural y funcional de CbpL 153

6.1	Caracterización de CbpL . . . . .	154
6.2	Obtención de la proteína soluble . . . . .	154
6.2.1	Clonación y expresión . . . . .	154
6.2.2	Purificación . . . . .	155
6.3	Cristalización . . . . .	157
6.4	Determinación estructural . . . . .	157
6.4.1	Huella peptídica de los cristales . . . . .	158
6.4.2	Determinación estructural de CbpL <sub>CBD</sub> . . . . .	158
6.5	Variantes del protocolo de purificación inicial . . . . .	160
6.5.1	Proteólisis de CbpL en función de la temperatura . . . . .	161
6.6	Análisis estructural de CbpL <sub>CBD</sub> . . . . .	162
6.6.1	Estructura general . . . . .	162
6.6.2	Sitios de unión a colina . . . . .	163
6.6.2.1	Sitios canónicos . . . . .	163
6.6.2.2	Sitios no canónicos . . . . .	163
6.6.3	Empaquetamiento cristalino . . . . .	164
6.6.4	Agitación térmica . . . . .	165
6.6.5	Otras moléculas y solvente . . . . .	165
6.7	Relación estructura-función . . . . .	166
6.7.1	Dominios Excalibur y Lipoprotein_Ltp . . . . .	168
6.7.2	Dominio CbpL <sub>CBD</sub> . . . . .	169

---

---

6.7.3	Localización y anclaje a la pared . . . . .	172
6.7.4	Estudios microbiológicos y fisiológicos relacionados con CbpL . . . . .	173
6.7.4.1	Caracterización fenotípica de células deficientes en CbpL . . . .	174
6.7.4.2	Modelos de infección en ratones y supervivencia en macrófagos .	174
<b>7.</b>	<b>Consideraciones finales</b>	<b>177</b>
7.1	MurA y el inicio de la síntesis de peptidoglicano . . . . .	177
7.2	DacB, degradación de pared y división celular . . . . .	179
7.3	CbpL, anclaje a pared y adhesión al hospedador . . . . .	181
<b>IV</b>	<b>CONCLUSIONES</b>	<b>183</b>
<b>V</b>	<b>BIBLIOGRAFÍA</b>	<b>187</b>
<b>VI</b>	<b>ARTÍCULOS DE INVESTIGACIÓN</b>	<b>211</b>





---

## ÍNDICE DE FIGURAS

---

1.1	<i>Streptococcus pneumoniae</i> . . . . .	4
1.2	Cápsula de <i>S. pneumoniae</i> . . . . .	5
1.3	Pared bacteriana . . . . .	7
1.4	Estructura general del peptidoglicano . . . . .	7
1.5	Esquema de la estructura del peptidoglicano . . . . .	8
1.6	Modelo de panal de abeja de peptidoglicano . . . . .	9
1.7	Modelo de ácido lipoteicoico . . . . .	10
1.8	Ruta de síntesis de peptidoglicano en <i>S. pneumoniae</i> . . . . .	11
1.9	División celular en <i>S. pneumoniae</i> . . . . .	13
1.10	Enzimas implicadas en la degradación de pared en <i>S. pneumoniae</i> . . . . .	14
1.11	Proteínas de superficie de <i>S. pneumoniae</i> . . . . .	15
1.12	Proteínas de unión a colina de <i>S. pneumoniae</i> . . . . .	17
3.1	Etapas básicas de la resolución estructural de proteínas mediante cristalografía de rayos X . . . . .	26
3.2	Diagrama de fases del proceso de cristalización . . . . .	32
3.3	Cristalización en presencia de ligandos . . . . .	34
3.4	Cristalización por difusión de vapor con aislamiento . . . . .	35
3.5	Cristalización por difusión de vapor sin aislamiento . . . . .	36
3.6	Defectos en el crecimiento de los cristales . . . . .	38
3.7	Esquema general del experimento de difracción . . . . .	39
3.8	Espectros de emisión de cobre y molibdeno y generador de ánodo rotatorio . . . . .	40
3.9	Instalación de radiación sincrotrón . . . . .	41
3.10	Montaje de un cristal en un difractómetro de rayos X o línea de radiación sincrotrón . . . . .	42
3.11	Detectores utilizados en cristalografía de rayos X . . . . .	44
3.12	Unidad asimétrica, celdilla unidad y cristal . . . . .	45
3.13	Red directa y red recíproca . . . . .	46
3.14	Ley de Bragg y planos de Miller . . . . .	46
3.15	Esfera de Ewald . . . . .	47
3.16	Patrón de difracción . . . . .	48
3.17	Defectos en los patrones de difracción . . . . .	49
3.18	Etapas del procesamiento de un patrón de difracción . . . . .	49

---

---

3.19	Método de reemplazo molecular . . . . .	54
3.20	Espectro de fluorescencia para la elección de la longitud de onda . . . . .	57
3.21	Ley de Friedel . . . . .	58
3.22	Indicadores del proceso de faseado experimental . . . . .	60
3.23	Mapas de densidad electrónica de uso frecuente . . . . .	61
3.24	Esquema del proceso de refinamiento . . . . .	62
3.25	Búsqueda de mínimos globales durante el refinamiento . . . . .	62
3.26	Comportamiento de los factores $R$ durante el refinamiento . . . . .	65
3.27	Tipos de macla en función del patrón de difracción . . . . .	67
3.28	Tipos de macla en función del número de dominios . . . . .	68
4.1	Secuencia de la proteína MurA1 de <i>S. pneumoniae</i> D39 . . . . .	74
4.2	Cristales de MurA1 <sub>D39</sub> . . . . .	76
4.3	Estructura general de MurA1 <sub>D39</sub> . . . . .	78
4.4	Simetría y empaquetamiento de MurA1 <sub>D39</sub> . . . . .	79
4.5	Agitación térmica de MurA1 <sub>D39</sub> . . . . .	80
4.6	Cristales de MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	81
4.7	Estructura general de MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	83
4.8	Simetría y empaquetamiento de MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	84
4.9	Agitación térmica de MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	85
4.10	Ligandos inespecíficos en MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	86
4.11	Alineamiento de secuencias de MurA1 <sub>D39</sub> y MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	88
4.12	Alineamiento estructural de MurA1 . . . . .	89
4.13	Alineamiento de los lazos flexibles de MurA en diferentes especies . . . . .	89
4.14	Esquema del mecanismo de reacción de MurA1 . . . . .	90
4.15	Interacciones implicadas en el cambio conformacional de MurA1 . . . . .	91
4.16	Residuos implicados en el reconocimiento de sustratos de MurA1 . . . . .	92
4.17	Comparación entre los sitios de unión de fosfomicina y citrato en MurA . . . . .	93
4.18	Efecto del citrato sódico en células de <i>S. pneumoniae</i> . . . . .	93
4.19	Mecanismo de reacción de MurA1 basado en la cisteína catalítica . . . . .	94
4.20	Mecanismo de reacción de MurA1 sin mediación de la cisteína . . . . .	95
4.21	Mecanismo de reacción de MurA1 con ruptura espontánea del THI . . . . .	95
4.22	Localización de Ala364Thr y organización del entorno . . . . .	96
5.1	Secuencia de la proteína DacB de <i>S. pneumoniae</i> TIGR4 . . . . .	100
5.2	Secuencia de la construcción DacB <sub>20–238</sub> . . . . .	101
5.3	Cristales de DacB <sub>20–238</sub> . . . . .	102
5.4	Espectro de fluorescencia teórico del zinc . . . . .	104
5.5	Estructura general de DacB <sub>20–238</sub> . . . . .	107
5.6	Simetría y empaquetamiento de DacB <sub>20–238</sub> . . . . .	107
5.7	Representación de factores de agitación térmica de DacB <sub>20–238</sub> . . . . .	108
5.8	Secuencia de la construcción DacB <sub>61–238</sub> . . . . .	109
5.9	Comparación de la secuencia de DacB y VanXYg . . . . .	110
5.10	Cristales de DacB <sub>61–238</sub> en gradiente de PEG 4000 . . . . .	111

---

---

5.11	Espectrometría de masas de DacB <sub>61–238</sub> . . . . .	113
5.12	Cristales de DacB <sub>61–238</sub> en gradiente de CaCl <sub>2</sub> . . . . .	114
5.13	Cristales de DacB <sub>61–238</sub> con variación de volúmenes . . . . .	114
5.14	Cristales de DacB <sub>61–238</sub> en presencia de EDTA . . . . .	115
5.15	Cristales de DacB <sub>61–238</sub> en presencia de MPD . . . . .	115
5.16	Cristales de DacB <sub>61–238</sub> obtenidos por sembrado . . . . .	116
5.17	Funciones de autorrotación de DacB <sub>61–238</sub> . . . . .	117
5.18	Sesgo en los mapas de densidad electrónica de DacB <sub>61–238</sub> . . . . .	119
5.19	Patrón de difracción de DacB <sub>61–238</sub> y espacio recíproco . . . . .	120
5.20	Defectos en los patrones de difracción de DacB <sub>61–238</sub> . . . . .	121
5.21	Espectros teóricos de absorción de los dispersores anómalos usados en el faseado de DacB <sub>61–238</sub> . . . . .	123
5.22	Espectro de fluorescencia de DacB <sub>61–238</sub> en torno al zinc . . . . .	125
5.23	Estructura general de DacB <sub>61–238</sub> . . . . .	125
5.24	Simetría y empaquetamiento de DacB <sub>61–238</sub> . . . . .	126
5.25	Representación de factores de agitación térmica de DacB <sub>61–238</sub> . . . . .	127
5.26	Secuencia de la construcción DacB <sub>49–238</sub> . . . . .	128
5.27	Cristales de DacB <sub>49–238</sub> . . . . .	128
5.28	Estructura general de DacB <sub>49–238</sub> . . . . .	131
5.29	Sitio de unión a zinc de DacB <sub>49–238</sub> . . . . .	131
5.30	Conformaciones abierta y cerrada de DacB <sub>49–238</sub> . . . . .	132
5.31	Simetría y empaquetamiento de DacB <sub>49–238</sub> . . . . .	133
5.32	Interacciones del extremo N-terminal de DacB <sub>49–238</sub> . . . . .	133
5.33	Comparación del empaquetamiento de DacB <sub>20–238</sub> y DacB <sub>49–238</sub> . . . . .	134
5.34	Representación de factores de agitación térmica de DacB <sub>49–238</sub> . . . . .	135
5.35	Interacciones de átomos de zinc en DacB <sub>49–238</sub> . . . . .	136
5.36	Comparación estructural entre DacB <sub>49–238</sub> y DacB <sub>61–238</sub> . . . . .	137
5.37	Superposición estructural de DacB con proteínas similares . . . . .	139
5.38	Estructura general de DacA . . . . .	140
5.39	Potencial electrostático de las superficies de DacA y DacB . . . . .	141
5.40	Sitios catalíticos de DacB y DacA . . . . .	142
5.41	Sitios de unión a sustrato en DacB y VanXYc . . . . .	143
5.42	Comparación de los lazos $\beta 5 - \alpha 4$ de las conformaciones abierta y cerrada de DacB . . . . .	143
5.43	Modelo computacional de L-Lys-D-Ala y residuos implicados en su reconocimiento . . . . .	144
5.44	Reconocimiento de sustrato en VanXYc y DacB . . . . .	145
5.45	Ensayos de expresión de DacA y DacB . . . . .	146
5.46	Efecto de DacA, DacB y Lgt en la morfología celular . . . . .	147
5.47	Análisis de los fragmentos de peptidoglicano de <i>s. pneumoniae</i> . . . . .	148
5.48	Efecto de DacA y DacB en la infección en ratones . . . . .	150
5.49	Efecto de DacA y DacB en la incorporación en macrófagos . . . . .	150
5.50	Efecto de DacA y DacB en la adherencia al hospedador . . . . .	151

---

---

6.1	Secuencia de la proteína CbpL de <i>S. pneumoniae</i> TIGR4 . . . . .	154
6.2	Similitud entre Q Sepharose® y la molécula de colina . . . . .	156
6.3	SDS-PAGE y cromatografía de exclusión molecular de CbpL . . . . .	157
6.4	Cristales obtenidos tras la purificación de CbpL . . . . .	158
6.5	Huella peptídica sobre los cristales de CbpL . . . . .	159
6.6	Efecto de la temperatura sobre la proteólisis de CbpL . . . . .	161
6.7	Estructura general de CbpL <sub>CBD</sub> . . . . .	162
6.8	Distribución de sitios canónicos y no canónicos de unión a colina en CbpL <sub>CBD</sub> . . . . .	163
6.9	Sitios canónicos de unión a colina en CbpL <sub>CBD</sub> . . . . .	164
6.10	Sitios no canónicos de unión a colina en CbpL <sub>CBD</sub> . . . . .	165
6.11	Simetría y empaquetamiento de CbpL <sub>CBD</sub> . . . . .	166
6.12	Relación entre el empaquetamiento de CbpL <sub>CBD</sub> y el último sitio de unión a colina . . . . .	167
6.13	Representación de factores de agitación térmica de CbpL <sub>CBD</sub> . . . . .	167
6.14	Densidad electrónica y factores de agitación térmica de una molécula de colina . . . . .	168
6.15	Comparación entre el dominio Excalibur y el motivo de unión a calcio de calmodulina . . . . .	168
6.16	Orientación teórica de CbpL en la superficie celular . . . . .	169
6.17	Comparación estructural de CbpL <sub>CBD</sub> con LytC y CbpF . . . . .	170
6.18	Sitios de unión a colina . . . . .	171
6.19	Distribución de sitios de unión a colina entre CBPs . . . . .	172
6.20	Distribución teórica de las uniones a ácidos teicoicos y lipoteicoicos . . . . .	173
6.21	Efecto de CbpL en la morfología celular . . . . .	174
6.22	Efecto de CbpL en la infección en ratones . . . . .	175
6.23	Efecto de CbpL en la incorporación en macrófagos . . . . .	175
7.1	Primera ruta de síntesis de peptidoglicano . . . . .	178
7.2	Modelo de degradación de peptidoglicano mediante DacA y DacB . . . . .	180
7.3	Ubicación de las CBPs en la superficie celular . . . . .	182

---

---

## ÍNDICE DE TABLAS

---

3.1	Relación de paquetes comerciales de cristalización . . . . .	37
3.2	Relación de líneas de radiación sincrotrón . . . . .	42
3.3	Indicadores estadísticos para el análisis de maclas . . . . .	69
4.1	Datos y estadísticas de difracción de MurA1 <sub>D39</sub> . . . . .	76
4.2	Estadísticas de refinamiento de MurA1 <sub>D39</sub> . . . . .	77
4.3	Datos y estadísticas de difracción de MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	82
4.4	Estadísticas de refinamiento de MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	82
5.1	Datos y estadísticas de difracción de DacB <sub>20–238</sub> . . . . .	103
5.2	Estadísticas relacionadas con el faseado experimental de DacB <sub>20–238</sub> . . . . .	105
5.3	Estadísticas de refinamiento de DacB <sub>20–238</sub> . . . . .	106
5.4	Datos y estadísticas de difracción preliminares de DacB <sub>61–238</sub> . . . . .	112
5.5	Leyes y fracciones de macla en DacB <sub>61–238</sub> . . . . .	118
5.6	Estadísticas de refinamiento de DacB <sub>61–238</sub> . . . . .	122
5.7	Estadísticas relacionadas con el faseado experimental de DacB <sub>49–238</sub> . . . . .	124
5.8	Datos y estadísticas de difracción de DacB <sub>49–238</sub> . . . . .	129
5.9	Estadísticas de refinamiento de DacB <sub>49–238</sub> . . . . .	130
6.1	Datos y estadísticas de difracción del dominio de unión a colina de CbpL . . .	159
6.2	Estadísticas de refinamiento de CbpL <sub>CBD</sub> . . . . .	160



---

## ABREVIATURAS

---

$CC^*$  : Coeficiente de correlación entre intensidades observadas y reales.

$CC_{1/2}$  : Coeficiente de correlación entre las reflexiones de dos subconjuntos.

$CC_{\text{all}}$  : Coeficiente de correlación para la estimación de la señal anómala considerando todas las reflexiones.

$CC_{\text{map}}$  : Coeficiente de correlación para la estimación de la subestructura atómica.

$CC_{\text{weak}}$  : Coeficiente de correlación para la estimación de la señal anómala considerando las reflexiones débiles.

$CC_{\text{work/free}}$  : Coeficiente de correlación entre el modelo y los datos de difracción.

$F_{hkl}$  : Factor de estructura de la reflexión de índices  $hkl$ .

$I_{hkl}$  : Intensidad de la reflexión de índices  $hkl$ .

$R_{\text{anom}}$  : Indicador de calidad de la señal anómala.

$R_{\text{merge}}$  : Factor de desacuerdo entre las diferentes medidas de una reflexión determinada.

$R_{\text{pim}}$  : Factor  $R_{\text{merge}}$  mejorado.

$R_{\text{work/free}}$  : Factor de desacuerdo entre el modelo y los datos de difracción.

$\alpha$  : Fracción de macla.

$\langle d''/\text{sig} \rangle$  : Señal anómala sobre ruido por capa de resolución.

$\langle |I|/\sigma(I) \rangle$  : Señal sobre ruido por capa de resolución.

$|F_{hkl}|$  : Módulo del factor de estructura.

$\varphi_{hkl}$  : Fase del factor de estructura.

$a, b, c, \alpha, \beta, \gamma$  : Parámetros de la celdilla unidad.

CBP : Proteína de unión a colina.

CCD : *Charged-Coupled Device*.

CFOM : Suma de  $CC_{\text{all}}$  y  $CC_{\text{weak}}$ .

---

CMI : Concentración Mínima Inhibitoria.

CMOS : *Complementary Metal-Oxide-Semiconductor*.

DLS : *Diamond Light Source*.

EP-UDP-GlcNAc : Enolpiruvil-UDP-*N*-acetilglucosamina.

ESRF : *European Synchrotron Radiation Facility*.

GlcNAc : *N*-acetilglucosamina.

HPAD : *Hybrid Pixel Array Detectors*.

IPTG : Isopropil- $\beta$ -D-1-tiogalactopiranosido.

MAD : Dispersión anómala múltiple.

MIR : Reemplazo isomorfo múltiple.

MPD : 2-metil-2,4-pentanodiol.

MurNAc : Ácido *N*-acetilmurámico.

NCS : Simetría no cristalográfica.

NCSP : Proteínas de superficie no-clásicas.

PBP : Proteína de unión a penicilina.

PDB : *Protein Data Bank*.

PEG : Polietilenglicol.

PEP : Fosfoenolpiruvato.

PMSF : Fluoruro de fenilmetilsulfonilo.

RFZ : Variable normalizada de la función de rotación.

SAD : Dispersión anómala simple.

SDS-PAGE : Electroforesis desnaturizante en gel de poliacrilamida.

SIR : Reemplazo isomorfo simple.

SLS : *Swiss Light Source*.

TFZ : Variable normalizada de la función de traslación.

THI : Intermediario tetrahédrico.

TLS : Traslación-Libración-Torsión.

UDP-GlcNAc : UDP-*N*-acetilglucosamina.

UDP-MurNAc : Ácido UDP-*N*-acetilmurámico.

---



---

## — ABSTRACT —

---

### **Characterization of peptidoglycan synthesis, degradation and binding proteins. Implications in virulence and antibiotic resistance mechanisms in *S. pneumoniae*.**

---

#### **• Introduction.**

*S. pneumoniae* is a Gram-positive pathogen responsible for high morbidity and mortality worldwide, mainly due to pneumonia, otitis, meningitis and sepsis. The high-risk groups are children, elderly persons and immunocompromised patients. Some of its most important virulence factors are the cell wall and associated proteins. Pneumococcal cell wall is composed of a thick layer of peptidoglycan that surrounds the cell surface, containing teichoic acids, lipoteichoic acids and different types of surface proteins.

Peptidoglycan is formed by long glycan chains of alterned units of *N*-acetylglucosamine and *N*-acetylmuramic acid and a stem peptide linked to the *N*-acetylmuramic acid moieties. These stem peptides are cross-linked through direct covalent bond or by the establishment of a dipeptide bridge. Teichoic acids (TA) are covalently attached to this network, while lipoteichoic acids (LTA) are attached to the lipid membrane and protrude to the surface of the cell through the peptidoglycan layer. In the pneumococcus, both TA and LTA are identical and show a unique feature, that is, the presence of molecules of phosphorylcholine that decorate pneumococcal surface. Finally, there is a plethora of surface proteins involved in essential processes such as peptidoglycan remodeling, host recognition and virulence mechanisms. Some surface proteins are attached to the lipid bilayer (lipoproteins), while others are attached to the choline moieties of the teichoic and lipoteichoic acids (the so-called Choline-Binding Proteins or CBPs). These CBPs are modular proteins that contain a functional domain and a choline-binding domain with different choline-binding sites, allowing interactions with the choline fragments of the cell wall through cation- $\pi$  interactions.

Peptidoglycan remodeling is a fundamental feature of any bacteria due to its involvement in cell growth, metabolites trafficking and virulence mechanisms. Cell wall remodeling requires two different pathways, namely the biosynthesis and the degradation. The biosynthetic pathway begins with the conversion of UDP-GlcNAc into UDP-MurNAc, through the action of MurA and MurB. Then, the stem peptide is added to the UDP-MurNAc, residue by residue, conforming a peptidoglycan moiety that is anchored to a lipid and translocated to the cell surface. Once there, other reactions occur in order to incorporate these fragments to the peptidoglycan layer, which is enlarged. The degradation of the peptidoglycan is carried out by both peptidases, that break down the bonds between the different residues from the stem peptide, and by lytic proteins, that cleave the glycan chains into monomers.

---

## • Objectives.

Because of the importance of the cell wall in virulence, three proteins from *S. pneumoniae* have been characterized in this work, namely, MurA1 (the first enzyme in the peptidoglycan synthesis pathway and also potentially involved in heteroresistance to fosfomycin), DacB (a lipoprotein that shows L,D-carboxypeptidase activity, cleaving the stem tetrapeptide during peptidoglycan degradation) and CbpL (a CBP involved in host-pathogen interactions during invasion). The proposed objectives for each protein follows:

### ■ MurA1:

- Crystallization and structural determination of MurA1<sub>D39</sub>.
- Crystallization and structural determination of MurA1<sub>Hungary<sup>19A</sup></sub>.
- Structural analysis of both proteins, in presence and absence of ligands of interest.
- Characterization of the heteroresistance to fosfomycin phenomenon by an structural approach.

### ■ DacB:

- Crystallization and structural determination of this lipoprotein.
- Structural characterization and analysis, in presence and absence of ligands of interest.
- Characterization of its mechanism of action and its relation to other proteins involved in the peptidoglycan degradation.

### ■ CbpL:

- Crystallization and structural determination of this choline-binding protein.
- Structural characterization and analysis.
- Functional characterization, implications in virulence and correlation with its three-dimensional structure.

## • Results.

The three-dimensional structures of MurA1<sub>D39</sub> and MurA1<sub>Hungary<sup>19A</sup></sub> have been determined at atomic resolution. They are formed by a single chain of 419 residues arranged in two globular domains linked by a double-stranded hinge. The catalytic site is located in the space between the two globular domains and a change of conformation is required for substrate recognition and reaction. MurA1<sub>D39</sub> crystallized in an open conformation, while MurA1<sub>Hungary<sup>19A</sup></sub> showed the closed conformation, due to a citrate molecule located in the catalytic site. This citrate acted as an inhibitor of the protein, allowing the comparison of the active and the inactive states. In addition, the heteroresistance to fosfomycin phenomenon has been analysed. Despite the mutation Ala364Thr (present in MurA1<sub>Hungary<sup>19A</sup></sub> but not in MurA1<sub>D39</sub>) was pointed out as the main cause for the change of heteroresistance phenotype, our work reveals that this structural difference is not likely involved in the heteroresistance mechanism.

---

Regarding lipoprotein DacB, its three-dimensional structure revealed the existence of a long flexible coil followed by the catalytic domain. The whole structure is covalently attached to the membrane. Its catalytic domain follows the  $\alpha + \beta$  globular folding, being structurally related to the vancomycin resistance D,D-carboxypeptidase VanXY. The catalytic pocket contains a mononuclear zinc binding site located in the middle of a large and fully exposed groove. Two different conformations were found presenting a different arrangement of the active site topology. The critical residues for catalysis and substrate specificity were identified. Loss-of-function of DacA and DacB altered the cell shape and peptide composition in *dac* mutants. In addition we demonstrate pathophysiological effects of disabled DacA or DacB activities. Real-time bioimaging of intranasal infected mice indicated a substantial attenuation of the infection in DacB-deficient mutants. In addition, uptake of DacA and DacB-deficient mutants by professional phagocytes was enhanced, while the adherence to lung epithelial cells was decreased, showing that both proteins could be optimal drug targets.

The choline-binding protein L, CbpL, is a modular protein composed by a calcium-binding domain (named Excalibur domain), followed by the choline-binding domain and by an exclusive domain involved in membrane recognition and protein stabilization. Despite production and purification were done with the full-length CbpL protein, crystallization was obtained only for the choline-binding domain. Intensive trials to avoid proteolysis during crystallization experiments failed. The three-dimensional structure of the choline-binding domain is composed of nine repeats that are arranged in four canonical choline-binding sites, whose sequence follows the consensus sequence of other CBPs, and four non-canonical sites that had not been found before in any other CBP. These novel non-canonical choline-binding sites are built by two consecutive repeats and present four aromatic residues (one tryptophan and three tyrosines) directly involved in phosphorylcholine recognition (versus the three aromatic residues, two tryptophans and one tyrosine, found in canonical sites). In addition, the alternance of canonical and non-canonical choline-binding sites along the choline-binding domain suggests a relation between its configuration and protein specificity and location over the cell surface. The coexistence of this choline-binding domain together with the membrane recognition domain allows CbpL to be attached to the cell wall and to the membrane at the same time, unique feature among CBPs. Finally, the infection capacity of CbpL-deficient mutants was analysed. Real-time bioimaging of intranasal infected mice showed a delay in the dissemination of the pathogen, while uptake by professional phagocytes showed an unexpected high survival ratio. Based on our structural and functional data, we proposed a model for *in vivo* functioning in which CbpL could be specifically recognizing some cell wall regions, orienting the Excalibur domain to allow adhesion functions in order to recruit/subvert some host proteins.

## • Conclusions.

The main conclusions obtained during this work can be summarized as follows:

1. MurA<sub>1D39</sub> is formed by two globular domains linked by a hinge that allow some structural flexibility, being the catalytic site located in between both domains. A flexible loop completes the catalytic machinery, being involved in substrate recognition and catalysis.

- 
2. MurA1<sub>Hungary</sub><sup>19A</sup> shows the same folding as MurA1<sub>D39</sub>. The only differential element is the Ala364Thr mutation in domain I.
  3. The location of a citrate molecule in the phosphoenolpyruvate recognition site in MurA1<sub>Hungary</sub><sup>19A</sup> results in a structural lock of its closed conformation, acting as a strong inhibitor of its activity.
  4. *S. pneumoniae* *murA1* gene is involved in heteroresistance to fosfomycin, a common phenotype among the most of the pneumococcal strains. Structural analysis between MurA1<sub>D39</sub> and MurA1<sub>Hungary</sub><sup>19A</sup> (non heteroresistant strain) showed that this phenotype is not due to any direct structural feature.
  5. *S. pneumoniae* TIGR4 DacB is a surface lipoprotein, showing L,D-carboxypeptidase activity, that is composed by a signal peptide, containing the *LipoBox* motif, followed by a long flexible region and the catalytic domain.
  6. The active site of DacB contains a catalytic zinc and a loop whose orientation defines the shape of the catalytic cavity and affects substrate recognition.
  7. Structural analysis of DacB allowed us to identify the residues involved in specific recognition of D-Ala–L-Lys terminal peptides of the stem peptide.
  8. DacA cleaves the terminal D-Ala residue of the stem pentapeptide resulting in a tetrapeptide. Sequential action of DacB cleaves the tetrapeptide, releasing the fourth D-Ala residue. This mechanism of action, together with the presence of DacB in the cell division septum, suggests the existence of a regulation of its activity through additional elements.
  9. DacA and DacB participate in the cell wall remodeling and their deletion ends in morphological alterations of the pneumococcal cells. Both enzymes present a radically different catalytic machinery and three-dimensional structures pointing to different *in vivo* functioning.
  10. DacA and DacB behave as fundamental virulence factors during the infective process and during the adhesion to host cells, speeding up their invasion and death.
  11. *S. pneumoniae* TIGR4 CbpL is a choline-binding protein formed by a calcium-binding domain, involved in adhesion to the host extracellular matrix, a choline-binding domain involved in the anchoring to the bacterial cell wall and an exclusive domain involved in the interaction with the lipid bilayer and the lipoteichoic acids.
  12. The choline-binding domain (CbpL<sub>CBD</sub>) is formed by eight choline-binding sites, four of them canonical and the other four non-canonical. Canonical choline-binding sites follow the consensus sequence and composition. The non-canonical choline-binding sites found in CbpL are exclusive of this protein and have not been previously observed in any other CBP.

- 
13. Some physiological assays have shown the influence of CbpL in the dissemination of the infection, a feature associated to the adhesine function. In addition, CbpL behaves as a key element for the degradation of the pathogen via immune system macrophages.



---

---

# INTRODUCCIÓN

---

---





---

# *Streptococcus pneumoniae*

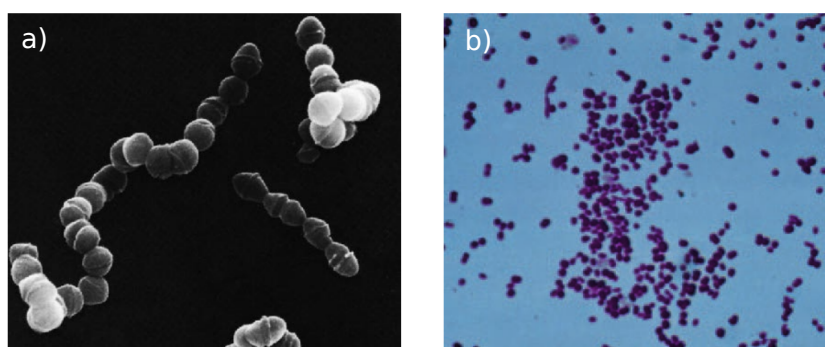
---

## Índice

1.1	Características generales . . . . .	4
1.2	Estructura de la pared celular . . . . .	6
1.3	Remodelado de la pared . . . . .	10
1.3.1	Biosíntesis . . . . .	10
1.3.2	División celular . . . . .	13
1.3.3	Degradación de peptidoglicano . . . . .	14
1.4	Proteínas de superficie . . . . .	14
1.4.1	Proteínas con motivo LPXTG . . . . .	15
1.4.2	Lipoproteínas . . . . .	16
1.4.3	Proteínas no-clásicas . . . . .	16
1.4.4	Proteínas de unión a colina . . . . .	16
1.5	Utilización de antibióticos y mecanismos de resistencia . . . . .	18
1.5.1	Fenómeno de heterorresistencia . . . . .	19

### 1.1. Características generales

*Streptococcus pneumoniae*, comúnmente denominado neumococo, fue aislado por primera vez por L. Pasteur y G. M. Sternberg en 1881, a partir de muestras de saliva de un niño con rabia (Pasteur, 1881; Sternberg, 1881) (Figura 1.1). Posteriormente, en 1886, A. Fraenkel demuestra que este microorganismo es un agente patógeno causante de la neumonía (Fraenkel, 1886). *S. pneumoniae* es una bacteria Gram positiva presente en la flora natural humana que se comporta como patógeno oportunista, causando una alta tasa de mortalidad (hasta un 30 %) especialmente en ancianos, niños y pacientes inmunocomprometidos (Musher, 1992). Los procesos infecciosos en los que está implicado provocan desde neumonía, otitis o sinusitis hasta afecciones más graves como meningitis o septicemia, pudiendo provocar la muerte del paciente. Según la Organización Mundial de la Salud (OMS), 1,1 millones de niños menores de cinco años mueren anualmente debido a la neumonía, lo que supone un 18 % de las defunciones infantiles a nivel mundial. De hecho, la neumonía es la primera causa de mortalidad infantil a nivel mundial, siendo la más común la originada por la infección de *S. pneumoniae*.

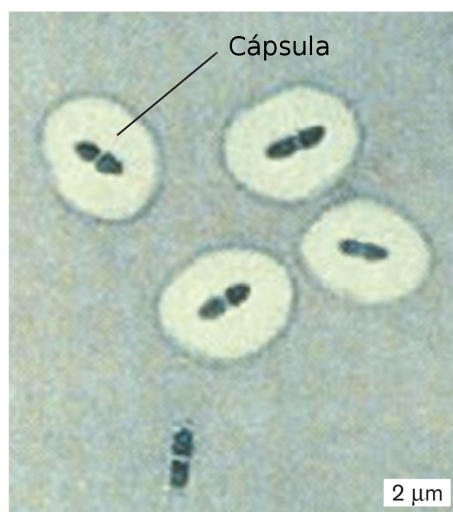


**Figura 1.1:** *Streptococcus pneumoniae*. a) Imagen de microscopía electrónica de barrido, a x33000. b) Imagen de microscopía óptica, a x900. Figura adaptada de Prescott *et al.*, 2002.

Su hábitat natural es la nasofaringe humana, llegando a esta región mediante una colonización producida los primeros días de vida de un neonato. No obstante, el mecanismo por el que alcanza su destino no ha sido descrito en su totalidad (Novak & Tuomanen, 1999). En general, el neumococo es capaz de permanecer en estado latente durante toda la vida, pero bajo diversas condiciones ambientales puede volverse virulento e iniciar una invasión (Cartwright, 2002). En concreto, los mayores problemas los causa en el tracto respiratorio inferior, acumulando fluidos y células del sistema inmune en los alveolos pulmonares. Además, puede extenderse desde el foco de la infección hacia huesos, oído medio y endocardio (Madigan *et al.*, 2004).

La sintomatología de una infección por neumococo es variable en función de la edad y del estado de salud del paciente. En personas sanas, suele iniciarse de forma repentina y se caracteriza por malestar y escalofríos intensos con temblores y fiebre, llegando hasta los 40 °C. Posteriormente se observan signos de dolor pleurítico, con posible sangrado. Las personas mayores no sufren una elevación tan brusca de la temperatura corporal. De hecho, en la mayoría de los casos el único signo de infección puede ser la pérdida de apetito y el deterioro del estado mental (Porth, 2007).

El principal factor de virulencia del neumococo es la cápsula (Figura 1.1). Este hecho fue determinado por F. Griffith en 1928, en un experimento que pasaría a formar parte de la historia de la microbiología. Dicho experimento consistió en la inoculación de cepas lisas (con cápsula) y cepas rugosas (sin cápsula) en ratones sanos, observando que los ratones tratados con la cepa lisa morían mientras que los infectados con la cepa sin cápsula sobrevivían (Griffith, 1928). La presencia de la cápsula retrasa o evita la fagocitosis por los macrófagos, disminuyendo notablemente la capacidad de la respuesta inmune contra el patógeno. Si bien posteriormente se han descubierto otras variantes, la variante capsulada es la más frecuente en las infecciones humanas por neumococo (Madigan *et al.*, 2004).



**Figura 1.2:** Cápsula de *S. pneumoniae*. Tinción con tinta china de células de *S. pneumoniae*. Figura adaptada de la galería del CDC (*Centers for Disease Control and Prevention*).

Una característica única y fundamental de *S. pneumoniae* es su dependencia de colina como elemento nutricional (Rane & Subbarow, 1940). La colina se incorpora en forma de fosforilcolina en los ácidos teicoicos y lipoteicoicos de la pared bacteriana (Tomasz, 1967; Bean & Tomasz, 1977; Behr *et al.*, 1992). La ausencia de colina en el medio provoca alteraciones estructurales reflejadas en una detención de la división celular y la formación de cadenas alargadas de células. Además, la sustitución de colina por análogos estructurales o la utilización de concentraciones elevadas de colina en el medio (más de 15  $\mu\text{g/mL}$ ) también dan lugar a las alteraciones anteriormente descritas. Otras alteraciones son posibles al sustituir la colina por etanolamina, como la tolerancia a ácidos  $\beta$ -lactámicos, la incapacidad de transformación y una mayor resistencia a infección por fagos. No obstante, estos efectos son reversibles bajo la adición de colina al medio (Tomasz, 1968; Tomasz, 1970). Por último, cabe destacar la existencia de cepas mutantes capaces de reproducirse y sobrevivir en ausencia de colina, como la R6Col<sup>-</sup> o la JY2190 (Severin *et al.*, 1997; Yother *et al.*, 1998). La morfología de *S. pneumoniae* no sólo se ve alterada por la ausencia de colina en el medio, sino que existen otros factores que pueden influir en el desarrollo de este patógeno. Como ejemplo, la ausencia de cationes divalentes influye en el proceso de división celular (Webb, 1951; Shankar & Bard, 1952).

Si bien la dependencia de *S. pneumoniae* por colina es única, existen otras especies del sistema respiratorio que exhiben colina en sus estructuras periféricas, como *S. mitis*, *S. oralis*, *Haemophilus influenzae*, *Neisseria meningitidis* o *Clostridium beijerinckii*. No obstante, la presencia de colina en la pared celular es un hecho poco habitual entre las bacterias (Tomasz, 1967; García *et al.*, 1998).

Como se ha indicado anteriormente, la envuelta celular de *S. pneumoniae* actúa como factor de virulencia, por lo que el estudio tanto de la cápsula como de la pared celular y las proteínas de superficie reporta información crucial en la lucha farmacológica contra este patógeno. Siendo el objetivo central de este trabajo el estudio de proteínas implicadas en el metabolismo y utilización de la pared bacteriana, se incluye una descripción detallada de este componente celular en las siguientes secciones.

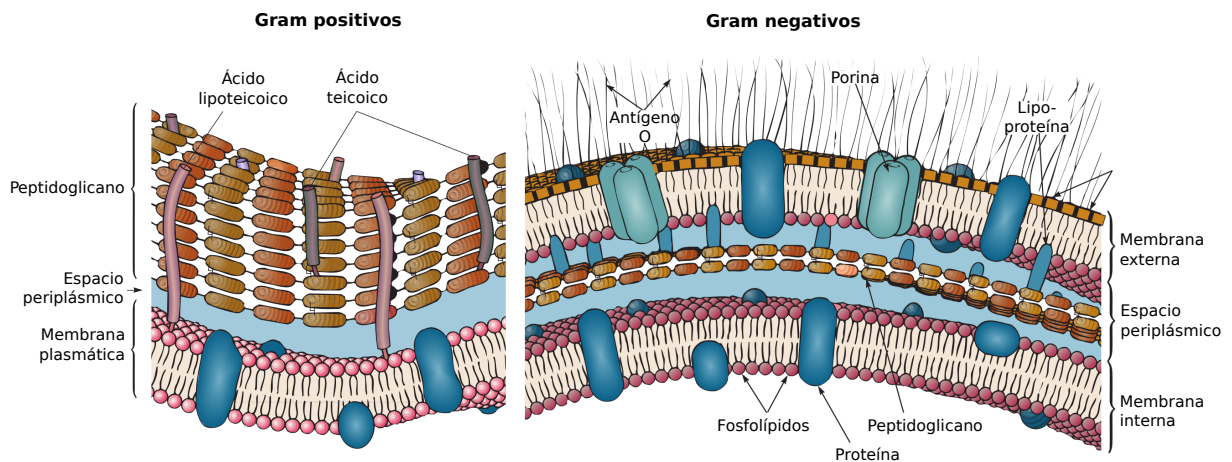
### 1.2. Estructura de la pared celular

La pared celular es la encargada de proteger a la bacteria de una posible lisis por diferencia de presión osmótica o por presión mecánica, además de modelar la morfología característica de cada especie. Su configuración también es importante desde el punto de vista de la permeabilidad, ya que debe permitir el transporte bidireccional entre el exterior y el interior celular de diversos compuestos implicados tanto en el metabolismo de la bacteria como en sus mecanismos de comunicación e infección (Madigan *et al.*, 2004)

Un aspecto especialmente interesante es el componente proteico de la pared bacteriana. Existen numerosas proteínas, ancladas a la pared, que ejercen muchas de las funciones descritas anteriormente (McCullers & Tuomanen, 2001). En el caso de *S. pneumoniae*, una serie de cambios metabólicos desembocan en alteraciones en la concentración y distribución de determinadas proteínas de superficie, repercutiendo en el comportamiento del patógeno durante la infección (Tuomanen *et al.*, 1985; Tuomanen *et al.*, 1987). El análisis del genoma de la cepa R6 de *S. pneumoniae* ha llevado a la identificación de más de 470 proteínas que incluyen un péptido señal para su exportación, bien para anclarse en la pared, bien para liberarse al medio extracelular (Hoskins *et al.*, 2001). Del genoma de la cepa TIGR4 se han obtenido alrededor de 70 proteínas expuestas a la superficie (Tettelin *et al.*, 2001).

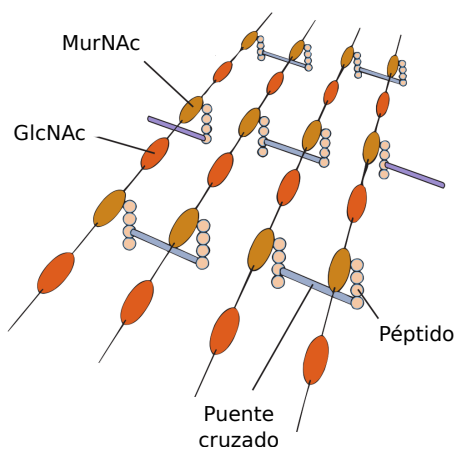
La composición de la pared varía entre especies. Si bien hay una clara dicotomía entre bacterias Gram positivas y Gram negativas (Figura 1.3), existen modificaciones menores que provocan que diferentes especies de un mismo grupo presenten paredes con una estructura específica. La pared de *S. pneumoniae* sigue la estructura general de especies Gram positivas, con una capa gruesa de peptidoglicano sobre la membrana lipídica y una serie de ácidos teicoicos y lipoteicoicos que atraviesan el peptidoglicano hasta contactar con la superficie o la cápsula. En total, el grosor de la pared oscila entre los 100 y los 400 Å (Madigan *et al.*, 2004).

El peptidoglicano consiste en una red de cadenas glicánicas formadas por la alternancia de unidades de *N*-acetilglucosamina (GlcNAc) y ácido *N*-acetilmurámico (MurNAc), unidas entre sí por fragmentos peptídicos (Figura 1.6). Las uniones entre GlcNAc y MurNAc se realizan a través de enlaces glicosídicos de tipo  $\beta 1 \rightarrow 4$ . Tras la incorporación de los azúcares a la



**Figura 1.3: Pared bacteriana.** La pared de bacterias Gram positivas presenta una única membrana lipídica y una capa gruesa de peptidoglicano, mientras que en organismos Gram negativos se observan dos membranas lipídicas, separadas por una capa delgada de peptidoglicano. Figura adaptada de Prescott *et al.*, 2002.

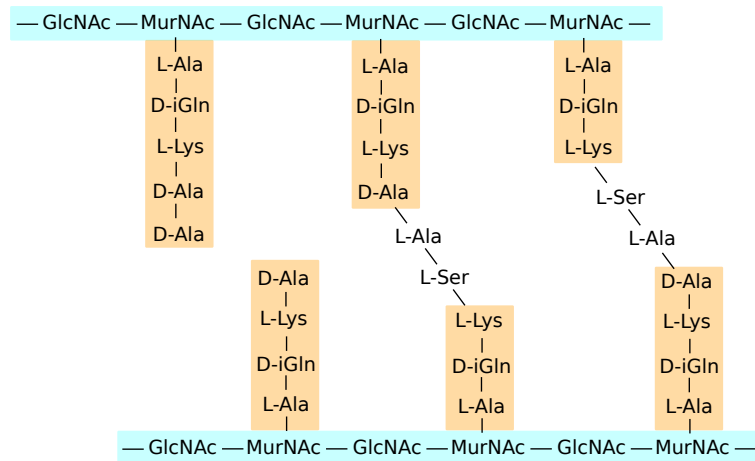
pared se inicia un proceso de deacetilación mediante la acción de la enzima PgdA, proceso que aporta mayor resistencia al ataque por lisozimas del hospedador. Se estima que los niveles de acetilación se sitúan en torno al 16 % para la GlcNAc y al 90 % para el MurNAc (Vollmer & Tomasz, 2000). Por su impacto en el proceso infeccioso, la deacetilación de la pared se considera factor de virulencia.



**Figura 1.4: Estructura general del peptidoglicano.** El peptidoglicano consta de cadenas glicánicas formadas por unidades de GlcNAc y MurNAc, unidas entre sí mediante fragmentos peptídicos. Figura adaptada de Prescott *et al.*, 2002.

Las cadenas glicánicas de GlcNAc y MurNAc contienen a su vez una serie de péptidos. En primer lugar, existe un fragmento peptídico que se une al MurNAc. En *S. pneumoniae* estos péptidos presentan una longitud máxima de cinco residuos (L-Ala-D-iGln-L-Lys-D-Ala-D-Ala), unidos por la L-Ala al grupo lactilo del MurNAc. No obstante, los péptidos de cinco residuos aparecen en menor concentración que los de tres (L-Ala-D-iGln-L-Lys) o cuatro (L-Ala-D-iGln-L-Lys-D-Ala) aminoácidos, mucho más frecuentes (Schleifer & Kandler, 1972) (Figura 1.5).

Los fragmentos peptídicos de diferentes cadenas glicánicas se unen entre sí para cerrar el entramado denominado peptidoglicano. En *S. pneumoniae* se producen dos tipos de uniones



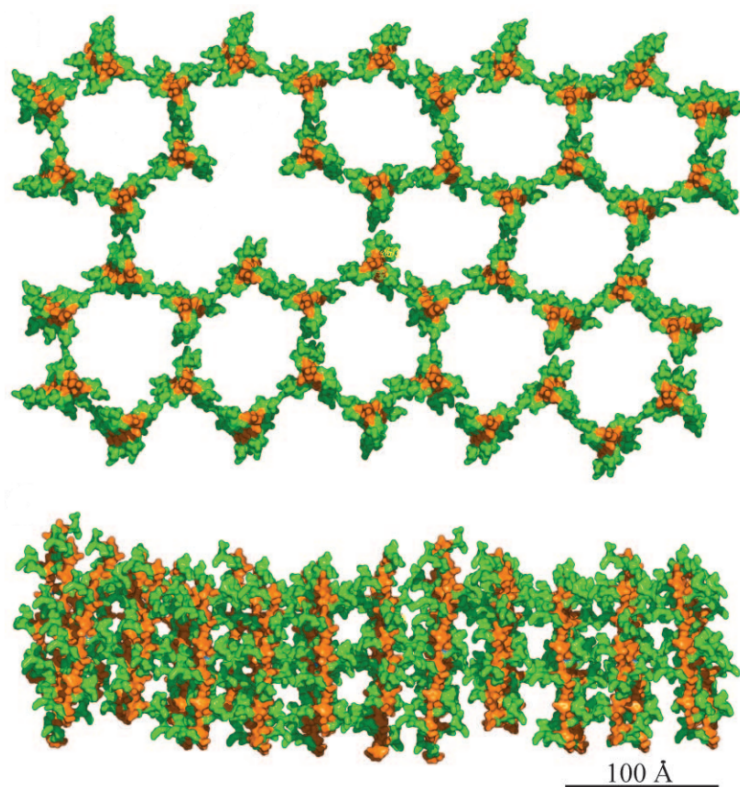
**Figura 1.5: Esquema de la estructura del peptidoglicano.** Las cadenas glicánicas se han sombreado en azul, mientras que los péptidos se han sombreado en naranja. Las uniones representadas son puentes de L-Ala-L-Ser, aunque también son posibles uniones de tipo L-Ala-L-Ala o directas entre la L-Lys y la D-Ala de los diferentes péptidos.

interpeptídicas. Las primeras consisten en un puente de L-Ala-L-Ser o L-Ala-L-Ala entre la última D-Ala del péptido de una cadena y el grupo  $\epsilon$ -amino de la L-Lys del péptido de otra cadena (Garcia-Bustos *et al.*, 1987) (Figura 1.5). El segundo tipo de unión consiste en un enlace directo entre la D-Ala terminal y la L-Lys, sin la mediación de un puente como en el caso anterior. Algunas cepas específicas presentan paredes basadas en pequeñas alteraciones de esta composición, lo que provoca cambios en la resistencia a determinados antibióticos (Fischer & Tomasz, 1985; Garcia-Bustos & Tomasz, 1990; Severin & Tomasz, 1996).

Las cadenas glicánicas se asocian entre sí formando estructuras helicoidales, quedando los fragmentos peptídicos orientados en diferentes direcciones. De esta forma se consigue que las cadenas glicánicas se asocien progresivamente formando una gran red (Madigan *et al.*, 2004). Además, existen poros que interrumpen el peptidoglicano en determinados puntos, permitiendo que las proteínas de superficie interaccionen con la membrana citoplasmática y con el exterior (Beveridge, 1981). El modelo más reciente, propuesto a partir de la estructura de un fragmento sintético de peptidoglicano (Meroueh *et al.*, 2006), sugiere que las cadenas glicánicas se sitúan de manera perpendicular a la membrana lipídica. De este modo, las cadenas peptídicas se organizan de forma paralela a la membrana, formando ángulos de  $120^\circ$  entre sí. Por tanto, la estructura resultante es de tipo panal de abeja, permitiendo la presencia de determinadas proteínas de superficie en los espacios vacíos resultantes (Figura 1.6).

Además del peptidoglicano anteriormente descrito, la pared celular contiene ácidos teicoicos y lipoteicoicos, descritos inicialmente como polisacárido-C y antígeno-F respectivamente (Tillet *et al.*, 1930; Goebel & Adams, 1943). Los ácidos teicoicos y lipoteicoicos de *S. pneumoniae* presentan la misma composición, un hecho único de esta especie (Fischer, 2000). Su estructura se basa en repeticiones de D-GalNAc(ChoP)-D-GalNAc(ChoP)-ribitol-5-P-6-D-Glc-AATGal, donde AATGal hace referencia a la 2-acetamido-4-amino-2,4,6-tridesoxi-D-galactosa y ChoP a unidades de fosforilcolina, unidas mediante enlace éster a la GalNAc. Estas unidades estructurales se unen entre sí mediante enlaces  $\alpha 1 \rightarrow 4$  entre la AATGal y la D-GalNAc de la siguiente unidad (Seo *et al.*, 2008; Gisch *et al.*, 2013) (Figura 1.7).



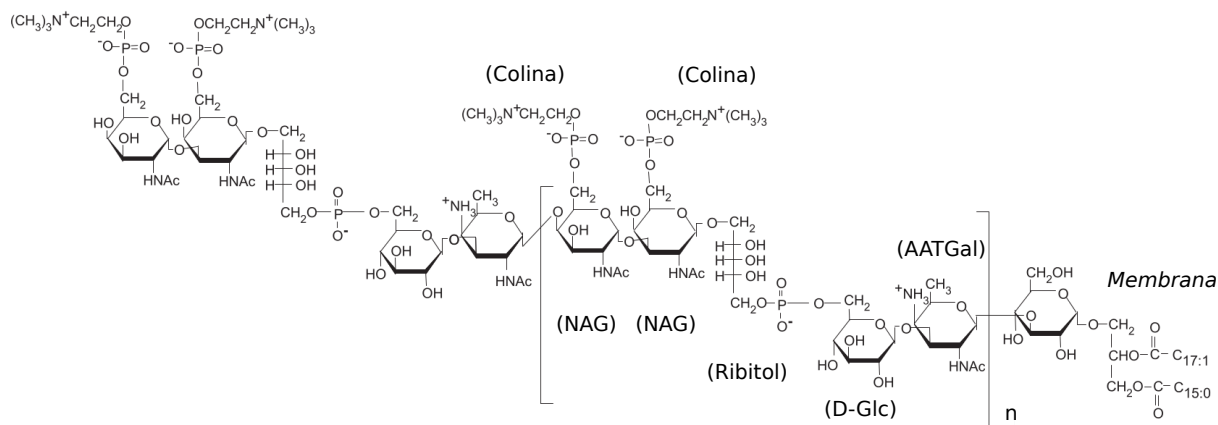


**Figura 1.6: Modelo de panal de abeja de peptidoglicano.** Estructura tridimensional de un fragmento sintético de peptidoglicano. Figura adaptada de Meroueh *et al.*, 2006.

Las longitudes de estas cadenas varían entre 2 y 8 repeticiones de cada unidad estructural. El porcentaje de fragmentos de fosforilcolina también es variable, dependiendo de la cepa, uniéndose preferentemente a la GalNAc cercana a la AATGal. Este azúcar no estándar aporta movilidad conformacional a la estructura y permite la asociación helicoidal de cadenas entre sí (Klein *et al.*, 1994; Fischer, 1997).

La diferencia principal entre ambos polímeros es que, mientras que los ácidos teicoicos se unen directamente al peptidoglicano, tanto en la cara interna como en la externa, los ácidos lipoteicoicos se anclan a la membrana citoplasmática (Skov-Sørensen *et al.*, 1988). La unión de los ácidos teicoicos a las cadenas glicánicas del peptidoglicano es de tipo covalente. Se estima que entre un 15 y un 19 % de las unidades de MurNAc del peptidoglicano unen cadenas de ácidos teicoicos, ya que han mostrado un grupo fosfato asociado al grupo hidroxilo anteriormente mencionado (Mosser & Tomasz, 1970; Yother *et al.*, 1998).

La unión de los ácidos lipoteicoicos a la membrana lipídica no se realiza por enlace directo entre las unidades estructurales y el fragmento de diacilglicerol, sino que media una D-glucosa (Figura 1.7). Este azúcar se une mediante enlace  $\alpha 1 \rightarrow 3$  al hidroxilo libre del diacilglicerol y mediante enlace  $\beta 1 \rightarrow 3$  a la AATGal de la unidad estructural previamente descrita (Glaser & Lindsay, 1974; Bracha *et al.*, 1978; Seo *et al.*, 2008; Gisch *et al.*, 2013). Los ácidos grasos asociados a los ácidos lipoteicoicos son tanto saturados como insaturados y de diferentes longitudes, si bien el más frecuente es el ácido palmítico (16:0), con una abundancia del 48 % (Behr *et al.*, 1992). Por último, en lo que a la distribución de moléculas de colina se refiere, se estima que un 85 % se



**Figura 1.7: Modelo de ácido lipoteicoico.** Se pueden observar las repeticiones de D-GalNAc(ChoP)-D-GalNAc(ChoP)-ribitol-5-P-6-D-Glc-AATGal y la unión a la membrana mediante una unidad de D-Glc. Figura adaptada de Seo *et al.*, 2008.

encuentra asociada a los ácidos teicoicos, estando el 15 % en los ácidos lipoteicoicos (Brundish & Baddiley, 1968).

Los ácidos teicoicos y lipoteicoicos se consideran factores de virulencia ya que activan la ruta alternativa del complemento y son altamente antigénicos, estimulando la formación de Il-1 en los monocitos y actuando como agentes quimiotácticos en determinadas inflamaciones (Tuomanen *et al.*, 1985; Fischer, 1988; Riesenfeld-Orn *et al.*, 1989). Por otro lado, las moléculas de colina participan activamente en el mecanismo de infección, siendo reconocidas tanto por los factores de activación plaquetaria de las células epiteliales del hospedador como por la proteína C reactiva humana, provocando la internalización del patógeno y la intensificación de la respuesta inmune respectivamente (Mold *et al.*, 1982; Cundell *et al.*, 1995; Hirschfield & Pepys, 2003). También participan en funciones fisiológicas propias de *S. pneumoniae* (López & García, 2004). Además, como se verá más adelante, las moléculas de colina no están implicadas únicamente en virulencia, sino que sirven de punto de anclaje para un tipo de proteínas de superficie denominadas proteínas de unión a colina o CBPs (*Choline Binding Proteins*; Swiatlo *et al.*, 2002).

### 1.3. Remodelado de la pared

Procesos como el crecimiento y la división celular requieren una remodelación constante de la pared bacteriana. Así pues, es fundamental disponer de un mecanismo tanto de biosíntesis como de degradación efectivo y localizado en el espacio y en el tiempo. El número de elementos implicados en estas rutas, ya sean de acción directa o de regulación, es muy elevado. En las siguientes secciones se ofrece una visión general de ambos mecanismos

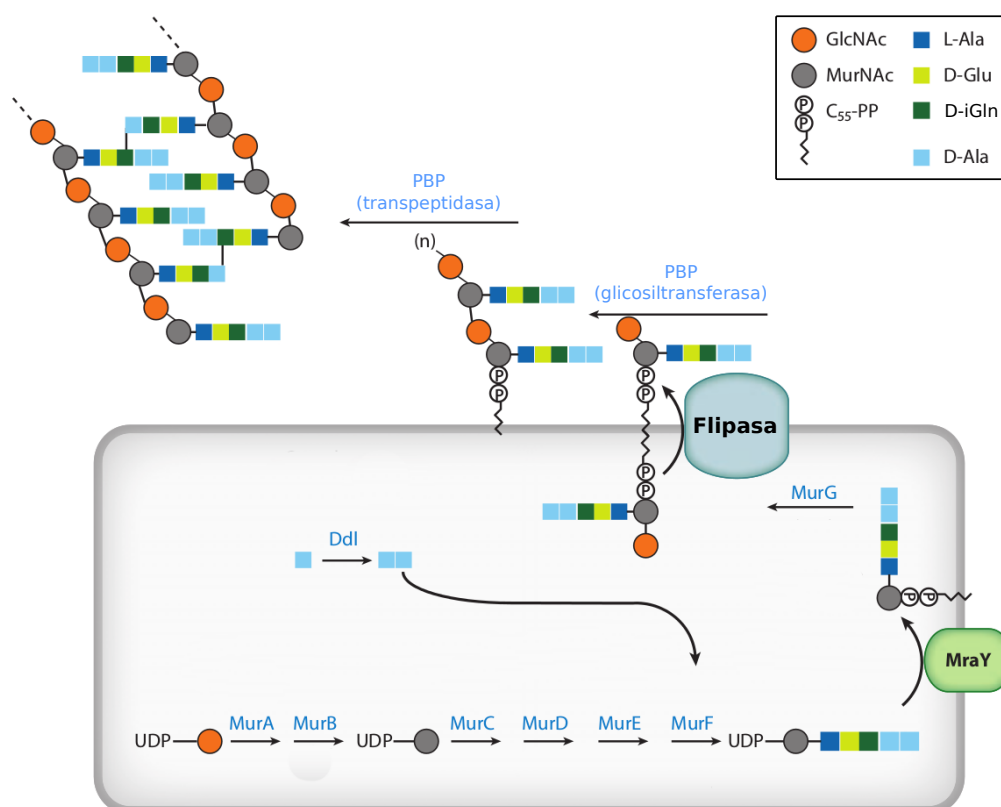
#### 1.3.1. Biosíntesis

La biosíntesis de pared en *S. pneumoniae* es bifásica (Laitinen & Tomasz, 1990). La fracción de pared naciente apenas establece enlaces cruzados entre sí, existiendo un porcentaje cercano al 70 % de péptidos no enlazados (mayoritariamente en su forma de tripéptido, L-Ala-D-iGln-L-Lys). Posteriormente, tras una etapa de maduración, el porcentaje de péptido libre descende



hasta un 40 %, lo que implica el establecimiento de enlaces cruzados y un aumento de la rigidez de la pared.

La ruta de biosíntesis de peptidoglicano implica a más de 20 proteínas. Generalmente se divide en tres etapas, a saber, la síntesis del precursor UDP-MurNAc–pentapéptido en el citoplasma celular, la formación del lípido II y su translocación hacia el espacio periplásmico y, por último, el crecimiento del peptidoglicano mediante el ensamblaje de las unidades anteriores (Figura 1.8). Como se ha mencionado anteriormente, la estructura resultante no se mantiene estática sino que sufre una serie de modificaciones por parte de enzimas ubicadas en la superficie celular.



**Figura 1.8: Ruta de síntesis de peptidoglicano en *S. pneumoniae*.** Figura adaptada de Lovering *et al.*, 2012.

La primera etapa de síntesis de pared de neumococo, que lleva al precursor del lípido II, es análoga a la de *E. coli*, perfectamente caracterizada (van Heijenoort, 1998). El proceso se inicia con la conversión de UDP-GlcNAc en UDP-MurNAc en dos pasos. En primer lugar, la proteína MurA, o UDP-GlcNAc enolpiruviltransferasa, cataliza la formación de enolpiruvil-UDP-GlcNAc (EP-UDP-GlcNAc). Posteriormente, este compuesto será procesado por la proteína MurB (UDP-GlcNAc enolpiruvilreductasa) para dar lugar a UDP-MurNAc (Sylvester *et al.*, 2001). Mientras que en especies Gram negativas sólo existe un único gen que codifique para la proteína MurA, en varias especies de Gram negativos (incluyendo *S. pneumoniae*) se encuentran dos genes, *murA1* y *murA2* que codifican para dos isoformas de la proteína MurA (Du *et al.*, 2000).

A continuación, las ligasas MurC, MurD, MurE y MurF actúan sobre el UDP-MurNAc para añadir, secuencialmente, los residuos L-Ala, D-Glu (en *S. pneumoniae* D-iGln), L-Lys y el dipéptido D-Ala-D-Ala, dando lugar finalmente al UDP-MurNAc-pentapéptido. La D-Ala y el D-iGln se generan a partir de sus respectivos L-aminoácidos mediante la acción de las racemasas Alr y MurI. El dipéptido D-Ala-D-Ala es sintetizado por la Ddl ligasa.

En este punto, en el que se dispone del precursor soluble UDP-MurNAc-pentapéptido, se inicia la segunda etapa de la biosíntesis de peptidoglicano. En primer lugar, la enzima MraY une el UDP-MurNAc-pentapéptido soluble al undecaprenilfosfato (C<sub>55</sub>-P), localizado en la cara citoplasmática de la membrana celular, liberando UMP. Al producto resultante se le denomina lípido I. A continuación, la proteína MurG añade UDP-GlcNAc libre del citoplasma al lípido I, dando lugar a C<sub>55</sub>-PP-MurNAc(pentapéptido)-GlcNAc, o lípido II.

EL lípido II sirve de sustrato a las transferasas MurM y MurN, que añaden los puentes de L-Ser-L-Ala o L-Ala-L-Ala para dar lugar a uniones cruzadas entre los péptidos (Filipe *et al.*, 2000; Filipe & Tomasz, 2000; Weber *et al.*, 2000). En concreto, MurM se encarga de unir el primer aminoácido del puente, mientras que MurN añade el segundo. La presencia de L-Ser o L-Ala en la primera posición depende del alelo de *murM* que se exprese.

La translocación del lípido II parece estar mediado por las proteínas FtsW y RodA, implicadas en los procesos de división y elongación celular (Pastoret *et al.*, 2004). Se cree que estas proteínas actúan como flipasas, dirigiendo la exportación del lípido II hacia el espacio extracelular, donde pasa a disposición de las PBPs (*Penicillin Binding Proteins*) (Matsushashi *et al.*, 1990). En concreto, la proteína FtsW se localiza junto a la transpeptidasa PBP3 (también denominada DacA, con actividad D,D-carboxipeptidasa), aunque no se ha demostrado que su asociación se relacione con su función (Morlot *et al.*, 2004).

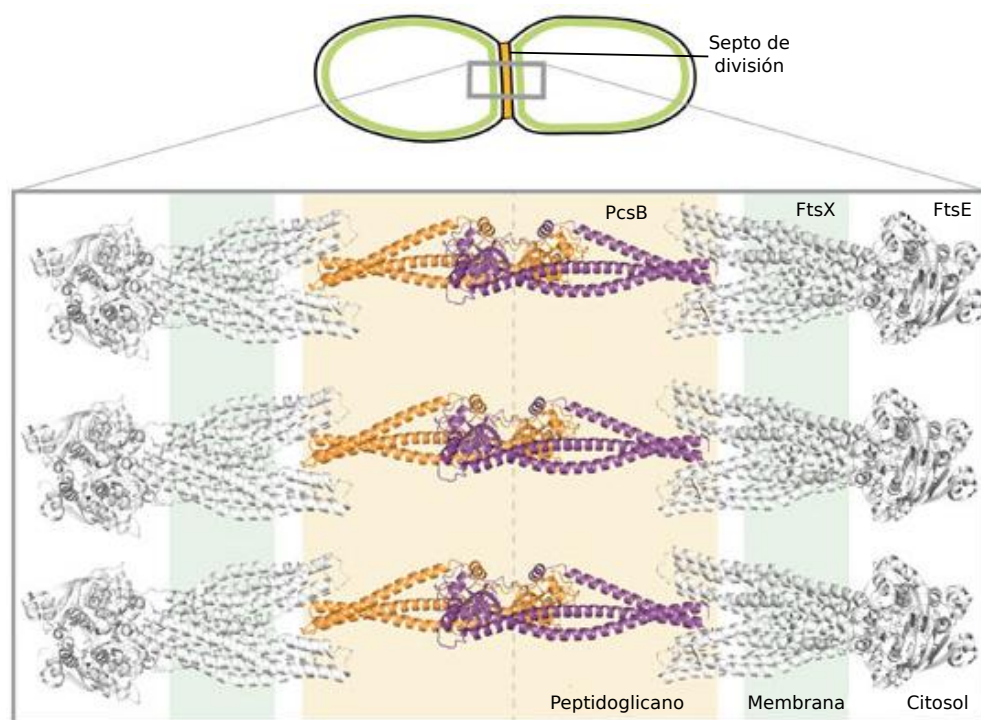
Por último, el lípido II se incorpora al peptidoglicano mediante la acción de transglicosilasas y transpeptidasas. En primer lugar, se forman polímeros de lípido II mediante reacciones de transglicosilación. Posteriormente, las transpeptidasas establecen los enlaces peptídicos cruzados, completando la estructura. Esta última reacción es llevada a cabo por las PBPs anteriormente mencionadas, tomando como sustrato un pentapéptido libre y enlazándolo con otro péptido mediante la liberación de la D-Ala terminal. Las PBPs se dividen en dos clases, a saber, la clase A, con proteínas que presentan actividad tanto transpeptidasa como transglicosilasa (como PBP1a, PBP1b y PBP2a), y la clase B, transpeptidasas monofuncionales (como PBP2b y PBP2x) (Goffin & Ghuysen, 1998).

En *S. pneumoniae*, los fragmentos nuevos de pared se incorporan en las zonas ecuatoriales de la bacteria, lo que lleva a la aparición de las zonas polares durante la división celular (Briles & Tomasz, 1970). Tras la división, cada célula presenta media pared antigua y media pared nueva (Morlot *et al.*, 2003). Esta distribución heterogénea conlleva que las diferentes proteínas implicadas en el proceso de biosíntesis de pared presenten localizaciones específicas sobre la superficie de la bacteria.

### 1.3.2. División celular

Existe una región genómica, denominada *dcw*, que contiene genes que codifican para proteínas implicadas en la biosíntesis de pared durante el proceso de división celular. Esta región aparece altamente conservada entre diferentes microorganismos de morfología similar, lo que sugiere una importancia crucial para la supervivencia de cada especie (Mingorance *et al.*, 2004). Algunas de los genes pertenecientes a esta región son *murD*, *murF*, *murG*, *mraY*, *pbp2b*, *pbp2x*, *ftsA* y *ftsZ* (Massidda *et al.*, 1998).

La división celular en *S. pneumoniae* contiene elementos análogos a los de otros microorganismos, pero el mecanismo no ha sido descrito en detalle (Massidda *et al.*, 2013). En general, el proceso se inicia con la formación de una estructura anular asociada a la cara interna de la membrana lipídica, compuesta por la proteína FtsZ, que rodea la bacteria. A esta estructura se le asocian posteriormente más de 10 proteínas distintas, dando lugar al denominado anillo Z (Goehring & Beckwith, 2005). A continuación se genera un septo de división entre las dos células, formado por una capa de peptidoglicano y algunas proteínas implicadas en el proceso de división (Figura 1.9). La proteína FtsA se localiza junto a FtsZ tanto en el septo de separación como en las zonas ecuatoriales de las células en división, interaccionando con ella (Lara *et al.*, 2005). Además de estas proteínas, en la zona de división también aparecen FtsB, FtsE, FtsL, FtsQ y FtsX (Noirclerc-Savoie *et al.*, 2005). Finalmente, el peptidoglicano del septo sufre una degradación parcial para permitir la separación de las células hijas, mediado por proteínas como PcsB (Figura 1.9) (Bartual *et al.*, 2014).

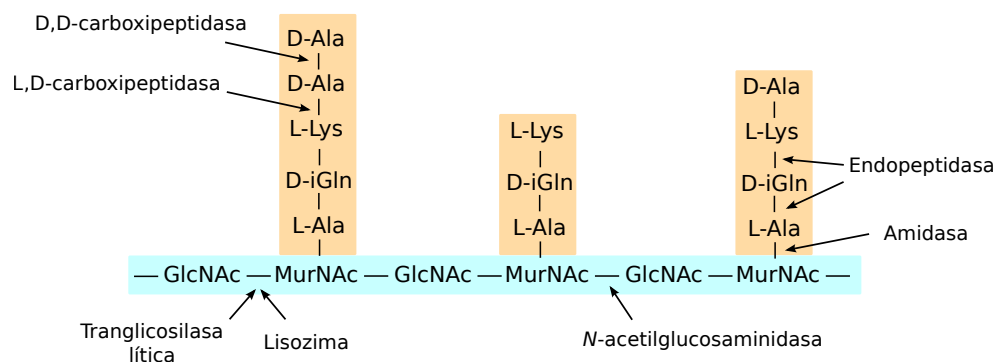


**Figura 1.9: División celular en *S. pneumoniae*.** Durante la división del neumococo se genera un septo constituido por una gruesa capa de peptidoglicano y diversas proteínas, como PcsB, que participa en su hidrólisis. Figura adaptada de Bartual *et al.*, 2014.

### 1.3.3. Degradación de peptidoglicano

Tan importante como la biosíntesis es la degradación de la pared, fundamental en los procesos de remodelación. Este fenómeno no es específico de neumococo, siendo las proteínas implicadas ubicuas en las bacterias. Además de participar en la rotura del septo de división, algunos de los mecanismos en los que se requiere una degradación del peptidoglicano son la regulación del alargamiento de la pared, el reciclaje de sus componentes o la autólisis (Shockman *et al.*, 1996).

Inicialmente, la presencia de tripéptidos en el peptidoglicano parecía indicar la existencia de D,D y L,D-carboxipeptidasas (García-Bustos *et al.*, 1987). Actualmente se conocen varias enzimas capaces de hidrolizar los enlaces covalentes (Figura 1.10). En primer lugar, la D-Ala–D-Ala carboxipeptidasa DacA (o PBP3) elimina la D-Ala terminal del pentapéptido. Posteriormente, la L-Lys–D-Ala carboxipeptidasa DacB elimina la D-Ala restante, dando lugar al tripéptido terminado en L-Lys (Barendt *et al.*, 2011). El péptido es separado de la MurNAc mediante la acción de LytA, una autolisina con actividad amidasa (Mellroth *et al.*, 2014). Las cadenas glicánicas restantes se degradan mediante la acción de la *N*-acetilglucosaminidasa LytB y la *N*-acetilmuramidasa LytC, que hidrolizan los distintos enlaces glicosídicos existentes entre la GlcNAc y el MurNAc (Gosink *et al.*, 2000). Finalmente, como se mencionó anteriormente, la proteína PgdA deacetila los fragmentos restantes (Vollmer & Tomasz, 2000; López & García, 2004).



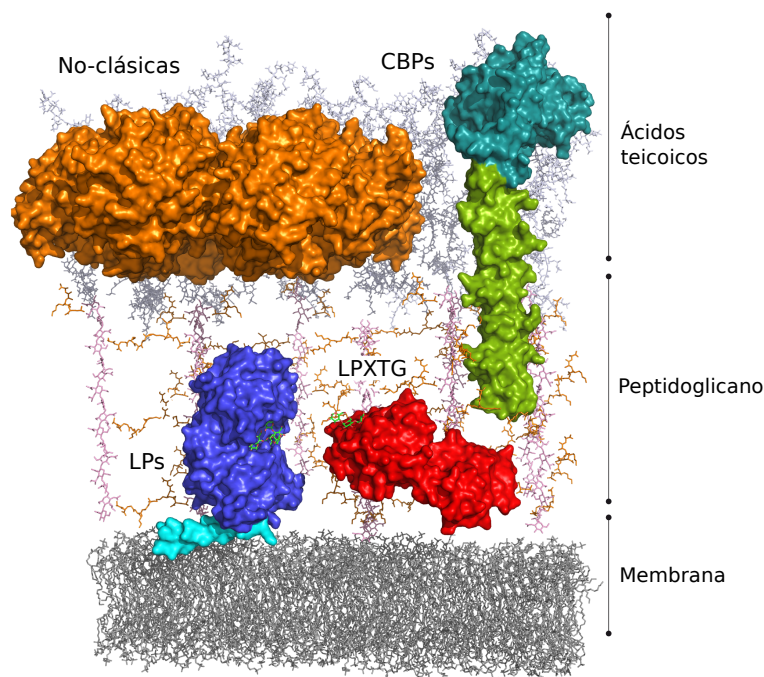
**Figura 1.10: Enzimas implicadas en la degradación de pared en *S. pneumoniae*.** Diversas proteínas participan en la rotura de los diferentes enlaces del peptidoglicano. Figura adaptada de Barendt *et al.*, 2011.

Existen otras proteínas que participan en los procesos de degradación de pared, tanto de forma directa como regulando el mecanismo en lugar y tiempo. Como sucedía con las proteínas de las últimas fases de la biosíntesis de pared, la mayoría de las proteínas implicadas en los mecanismos de degradación se ubican en la superficie de la bacteria, utilizando diferentes tipos de anclajes.

## 1.4. Proteínas de superficie

Si bien la pared bacteriana es un elemento de importancia vital para la célula, también lo son, como se ha indicado anteriormente, las proteínas de superficie. Debido a su implicación directa en las interacciones hospedador-patógeno, las proteínas de superficie se consideran factores

de virulencia (Mitchell, 2003; Hammerschmidt, 2006). Existen cuatro grupos de proteínas de superficie en *S. pneumoniae*, a saber, proteínas con el motivo LPXTG (o simplemente proteínas LPXTG), lipoproteínas, proteínas de unión a colina (CBPs) y proteínas de superficie no-clásicas (NCSPs) (Pérez-Dorado *et al.*, 2012) (Figura 1.11). Como característica diferencial, las NCSPs se encuentran en la envuelta celular, sin anclarse directamente a la membrana o al peptidoglicano. En cambio, el resto de proteínas de superficie presentan algún tipo de asociación a la membrana o a la pared bacteriana (Bergmann & Hammerschmidt, 2006)



**Figura 1.11: Proteínas de superficie de *S. pneumoniae*.** Las proteínas de superficie de neumococo se dividen en cuatro clases, a saber, proteínas con motivo LPXTG, lipoproteínas (LPs), proteínas de unión a colina (CBPs) y proteínas de superficie no-clásicas. Figura adaptada de Pérez-Dorado *et al.*, 2012.

#### 1.4.1. Proteínas con motivo LPXTG

Las proteínas con motivo LPXTG se anclan covalentemente a la pared bacteriana. Utilizan dos péptidos señal, en los extremos N y C-terminales, para alcanzar su ubicación final (Dramsi *et al.*, 2008; Löffling *et al.*, 2011). El péptido señal del extremo C-terminal contiene la secuencia LPXTG, seguida de una región hidrofóbica y una cola cargada positivamente, con la que mantiene la unión a la membrana tras su translocación. El enlace covalente implicado en esta unión es producido por una sortasa de clase A (SrtA) (Mazmanian *et al.*, 2000; Dramsi *et al.*, 2008).

El análisis de los genomas de *S. pneumoniae* TIGR4 y R6 ha revelado la existencia de 18 y 15 proteínas de tipo LPXTG respectivamente (Tettelin *et al.*, 2001; Frolet *et al.*, 2010; Löffling *et al.*, 2011). En general, este tipo de proteínas son modulares con funciones relacionadas con la adhesión, fundamentales en el proceso de invasión. Otro subgrupo presenta funciones enzimáticas de reconocimiento de células del hospedador. Entre estas últimas se encuentran glicosilhidrolasas (como StrH, NanA o EndoD), proteasas (ZmpA, ZmpB, ZmpC y PtrA) y

una con actividad polisacárido liasa (SpnHL), actuando principalmente sobre glicoproteínas y polímeros de azúcares implicados en diversos procesos biológicos vitales para el hospedador (Garbe & Collin, 2012).

### 1.4.2. Lipoproteínas

Las lipoproteínas presentan una modificación del extremo N-terminal basada en la adición de un grupo *N*-acildiacilgliceril que media el anclaje de la proteína a la membrana lipídica. La ruta de síntesis de este tipo de proteínas es común en gran número de especies de bacterias, lo que revela su importancia para la supervivencia de la célula.

Estas proteínas contienen un péptido señal seguido de un motivo denominado *LipoBox*. Este motivo, de secuencia consenso [LVI][ASTVI][GAS][C], es reconocido antes de añadir la modificación lipídica a la proteína. En concreto, la cisteína es la que realiza el enlace covalente con la fracción diacilgliceril que media la unión a membrana (Oudega *et al.*, 1993; Kovacs-Simon *et al.*, 2011).

Se ha predicho la existencia de hasta 46 lipoproteínas en *S. pneumoniae* (Hoskins *et al.*, 2001; Tettelin *et al.*, 2001; Babu *et al.*, 2006), la inmensa mayoría implicadas en la adecuación de la bacteria al medio en el que se encuentra. En concreto, más de la mitad son transportadores de tipo ABC, implicados en la unión y transformación de sustratos esenciales para el neumococo. El resto de lipoproteínas aparecen involucradas en procesos de plegamiento de proteínas y activación de otras proteínas de superficie (Babu *et al.*, 2006; Manzano *et al.*, 2008).

### 1.4.3. Proteínas no-clásicas

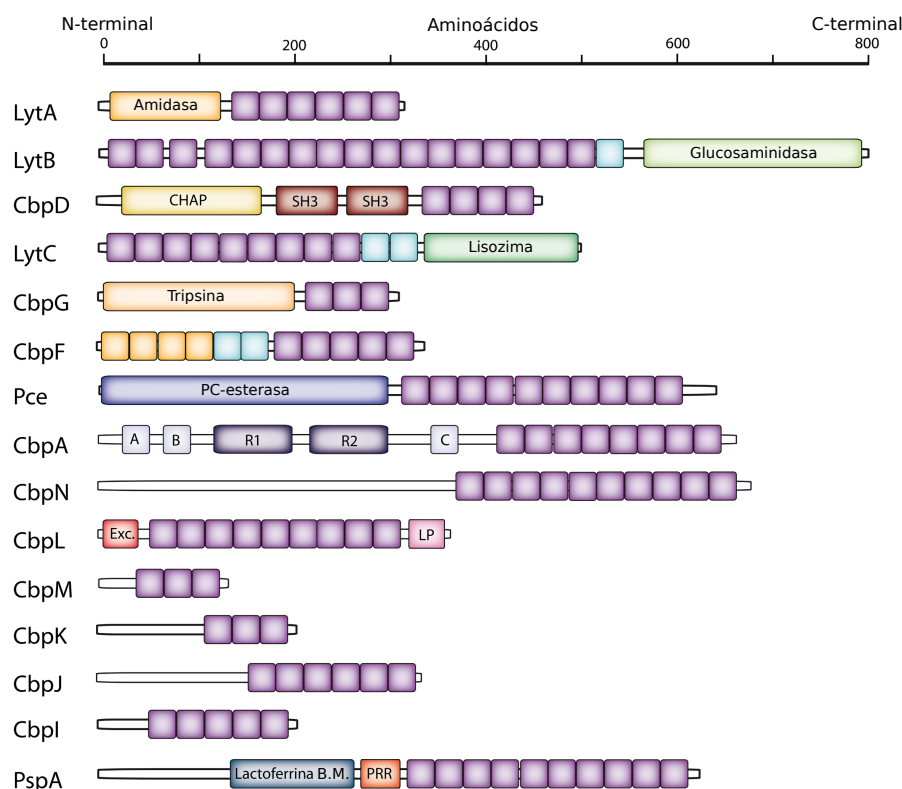
Estas proteínas, específicas de *S. pneumoniae* carecen de anclaje a membrana y péptidos señal, pese a localizarse en su superficie. Curiosamente, desempeñan diversas funciones en el citoplasma celular hasta que son exportadas a la superficie, por mecanismos aún desconocidos. Una vez ubicadas en la superficie, realizan funciones de adhesión al hospedador. Algunos ejemplos son PavA, que une fibronectina, o la fosfoglicerato quinasa (PGK), que une plasminógeno (Holmes *et al.*, 2001; Fulde *et al.*, 2014).

Hasta la fecha sólo se han identificado 6 proteínas de superficie no-clásicas, aunque la ausencia de péptidos señal o secuencias consenso dificulta su detección a partir de secuencia. De hecho, se cree que el número de este tipo de proteínas es bastante elevado.

### 1.4.4. Proteínas de unión a colina

Las proteínas de unión a colina, o CBPs, presentan una organización modular que contiene, al menos, el dominio de unión a colina y un dominio encargado de realizar alguna función biológica (Figura 1.12). El dominio de unión a colina interactúa con los fragmentos de colina de los ácidos teicoicos y lipoteicoicos, efectuando así el anclaje de toda la proteína a la pared.

El dominio de unión a colina se sitúa generalmente en el extremo C-terminal, salvo en el caso de LytB y LytC, que lo incluyen en el extremo N-terminal (Gosink *et al.*, 2000; Pérez-Dorado *et al.*, 2010). Su estructura se compone de una serie de repeticiones estructurales de unos 20



**Figura 1.12: Proteínas de unión a colina de *S. pneumoniae*.** Se muestran los dominios de los que consta cada CBP. Los recuadros morados hacen referencia a las repeticiones de unión a colina. Figura adaptada de Pérez-Dorado *et al.*, 2012.

residuos que dan lugar a los diferentes sitios de unión a colina. Estos sitios están formados por diferentes configuraciones de anillos aromáticos que estabilizan las moléculas de colina de los ácidos teicoicos y lipoteicoicos a través de interacciones catión- $\pi$  (García *et al.*, 1998).

Algunas proteínas de esta familia están implicadas en adhesión y virulencia, mientras que otras presentan actividad mureín hidrolasa. Además, se estima que las CBPs podrían estar implicadas en el establecimiento de interacciones no específicas con el hospedador, estabilizando electrostáticamente la cápsula por la carga negativa que aportan a la superficie celular (Swiatlo *et al.*, 2002; Hammerschmidt, 2006; Frolet *et al.*, 2010). Algunos ejemplos interesantes de CBPs son Pce (o CbpE), que participa en funciones de patogénesis al eliminar la capa existente de fosforilcolina y también el segundo mensajero PAF (Hermoso *et al.*, 2005), PspA, que inhibe el sistema de complemento del hospedador (Tu *et al.*, 1999) y participa en la evasión de la fagocitosis por macrófagos (Neeleman *et al.*, 1999), CbpA, que interacciona con citoquinas humanas (Rosenow *et al.*, 1997) y se une a la IgA humana y a la proteína C3 del sistema de complemento (Hammerschmidt *et al.*, 1997) o CbpG, que participa tanto en adhesión y colonización como en la sepsis, dependiendo de si se expresa el dominio de unión a colina junto al dominio catalítico o si únicamente se expresa el dominio catalítico (Mann *et al.*, 2006).

Otros ejemplos importantes de CBPs en *S. pneumoniae* son las autolisinas. Estas proteínas degradan la pared propia de neumococo para facilitar determinadas tareas como la división celular o el establecimiento de células competentes con capacidad de transformación. En concreto, las proteínas LytA y LytC liberan al medio componentes de la pared celular que



actúan como factores de virulencia (Berry *et al.*, 1989; Pérez-Dorado *et al.*, 2010), mientras que LytB, implicada en los mecanismos de separación de las células hijas, favorece el aumento de la carga microbiana y la diseminación durante una infección (de las Rivas *et al.*, 2002).

No sólo *S. pneumoniae* expresa proteínas de unión a colina, sino que algunas endolisinas codificadas por fagos de neumococo también presentan un dominio de unión a colina. Estas enzimas líticas facilitan la salida de los fagos tras el ciclo lisogénico (Lopez *et al.*, 1982). Todos los fagos de *S. pneumoniae*, salvo Cp-7, expresan endolisinas de tipo CBP con el módulo catalítico en el extremo N-terminal (García *et al.*, 1990).

### 1.5. Utilización de antibióticos y mecanismos de resistencia

El tratamiento contra *S. pneumoniae*, como en muchas otras enfermedades infecciosas, consiste en la utilización de antibióticos. Inicialmente, el neumococo era uniformemente sensible a penicilina aunque la aparición de resistencias ha motivado la búsqueda de nuevos antibióticos y nuevas estrategias terapéuticas (Hakenbeck *et al.*, 1999; Morand & Mühlemann, 2007). Como alternativa a la penicilina se ha utilizado eritromicina, cefalosporina, fluoroquinolona, ceftriaxona, cefotaxima o vancomicina (Madigan *et al.*, 2004).

Si bien las resistencias a antibióticos pueden generarse de forma natural, el abuso en su utilización durante las últimas décadas ha disparado la aparición de cepas resistentes a un rango amplio de compuestos. Un ejemplo extremo es el caso del antibiótico ceftarolina contra la proteína PBP2a de *Staphylococcus aureus*, comercializado en el año 2010, para el que se observaron las primeras cepas resistentes en el año 2011.

Los antibióticos  $\beta$ -lactámicos, como la penicilina, se unen irreversiblemente a las PBPs, impidiendo la formación de los enlaces cruzados entre los péptidos del peptidoglicano. Así pues este antibiótico debilita considerablemente la pared, evitando el crecimiento y diseminación del patógeno. La mayoría de las cepas resistentes a penicilina presentan una pared celular de composición diferente a la habitual, observándose un mayor número de enlaces cruzados entre péptidos (García-Bustos & Tomasz, 1990; Severin *et al.*, 1995). Se cree que este fenómeno se debe a que las PBPs de las cepas resistentes, además de disminuir su afinidad por penicilina, han modificado ligeramente su especificidad de sustrato (Dessen *et al.*, 2001). No obstante, se han observado cepas resistentes a penicilina con una estructura de pared estándar, por lo que la hipótesis anterior puede que sólo se refleje en algunas PBPs de cepas determinadas (Severin *et al.*, 1996).

La delección de *murMN* conlleva una pérdida de resistencia a penicilina, lo que sugiere que la presencia de enlaces cruzados entre los péptidos del peptidoglicano es necesaria para la aparición de la resistencia a este antibiótico (Filipe & Tomasz, 2000; Weber *et al.*, 2000). Estas cepas mutantes no sólo son sensibles a  $\beta$ -lactamasas, sino que también sufren el ataque de fosfomicina, cicloserina, nisina y vancomicina.

Si bien se siguen desarrollando antibióticos que actúan contra determinadas proteínas de *S. pneumoniae*, existen estrategias orientadas a la utilización de vacunas basadas en los



determinantes antigénicos de la cápsula y de las proteínas de superficie. En concreto, se ha desarrollado una vacuna polivalente compuesta por antígenos de 23 tipos de polisacáridos capsulares. Esta vacuna no se aplica a menores de 2 años, ya que su sistema inmunológico aún no ha madurado lo suficiente como para que se produzca una respuesta efectiva. Se recomienda la vacunación con este sistema en personas con enfermedades cardiovasculares o pulmonares crónicas, especialmente si son mayores de 65 años o inmunocomprometidas. Con una dosis se suele obtener inmunidad de por vida, aunque en determinados casos conviene administrar una segunda dosis. Posteriormente se desarrolló otra vacuna de polisacáridos y proteínas de superficie heptavalente, autorizada para su uso en niños menores de 2 años (Porth, 2007). Otra estrategia alternativa es la aplicación de los denominados *enzibióticos*, basados en proteínas líticas de fagos que actúan sobre *S. pneumoniae* (Hermoso *et al.*, 2007).

### 1.5.1. Fenómeno de heterorresistencia

Además de las resistencias observadas habitualmente frente a diferentes antibióticos, existe un fenómeno de heterorresistencia de especial relevancia en *S. pneumoniae*. La heterorresistencia a determinado antibiótico se define como la capacidad que presenta una población aislada para producir subpoblaciones en presencia de concentraciones de dicho antibiótico superiores a las de su concentración mínima inhibitoria (CMI) (Morand & Mühlemann, 2007). En otras palabras, las cepas heterorresistentes crecen por encima de las concentraciones efectivas de inhibición de determinados antibióticos, sobreviviendo a los tratamientos habituales y falseando las pruebas clínicas de detección de resistencias (Falagas *et al.*, 2008; Zhang *et al.*, 2014).

Este fenómeno ha sido estudiado en varios microorganismos y contra varios antibióticos. Actualmente se siguen realizando estudios para detectarlo y bloquearlo (Morand & Mühlemann, 2007; Deresinski, 2013; Folkvardsen *et al.*, 2013; Bianco *et al.*, 2014; Zavascki *et al.*, 2014; Nodari *et al.*, 2015). No obstante, los mecanismos de acción aún no han podido ser descritos en detalle, o incluso se desconocen por completo. Además, analizar las características genotípicas de determinadas subpoblaciones no resulta tarea sencilla, debido a la heterogeneidad existente entre los miles de millones de células que las componen (Avery, 2006).

Trabajos previos realizados en el grupo de la Prof. Kathrin Mühlemann y la Dra. Lucy J. Hathaway (Instituto de Enfermedades Infecciosas, Facultad de Medicina, Universidad de Berna, Suiza) han demostrado la existencia de heterorresistencia en *S. pneumoniae* a penicilina (Morand & Mühlemann, 2007; Engel *et al.*, 2014). Posteriormente se analizó la heterorresistencia de *S. pneumoniae* a fosfomicina, antibiótico de amplio espectro que se introduce en el centro activo de la proteína MurA y bloquea la síntesis de peptidoglicano. Mientras que la mayoría de cepas de *S. pneumoniae* presentan heterorresistencia a fosfomicina (como es el caso de las cepas D39, Spain<sup>9V</sup> o Finland<sup>6B</sup>), la cepa clínica Hungary<sup>19A</sup> carece de dicho fenotipo. En concreto, se ha observado la presencia de una mutación por sustitución (Ala364Thr) en el gen *murA1* de la cepa Hungary<sup>19A</sup> respecto a las cepas heterorresistentes. Así pues, un estudio comparado entre la proteína MurA de la cepa Hungary<sup>19A</sup> y la misma proteína de alguna cepa heterorresistente podría revelar indicios sobre el origen y la causa de este fenómeno.



---

## OBJETIVOS

---

Con la finalidad de ampliar el conocimiento sobre la biosíntesis y el remodelado de la pared bacteriana de *S. pneumoniae*, se han elegido tres proteínas implicadas en diferentes aspectos relacionados con el remodelado de peptidoglicano y su utilización. En primer lugar, MurA1, proteína citosólica encargada del primer paso de síntesis de peptidoglicano e implicada en el mecanismo de heterorresistencia a fosfomicina. En segundo lugar, DacB, lipoproteína de membrana implicada en su degradación y, por último, CbpL, una proteína de unión a colina, con posible comportamiento de lipoproteína, que utiliza los ácidos teicoicos y lipoteicoicos del peptidoglicano como soporte para realizar funciones de adhesión a las células del hospedador.

A continuación se enumeran los objetivos propuestos en este trabajo para cada una de las proteínas anteriormente descritas:

■ **MurA1:**

- Cristalización y determinación estructural de la proteína MurA1<sub>D39</sub>.
- Cristalización y determinación estructural de la proteína MurA1<sub>Hungary<sup>19A</sup></sub>.
- Análisis estructural de ambas proteínas, en presencia y ausencia de ligandos de interés.
- Caracterización del fenómeno de heterorresistencia a fosfomicina mediante una aproximación estructural.

■ **DacB:**

- Cristalización y determinación estructural.
- Análisis estructural, en presencia y ausencia de ligandos de interés.
- Caracterización del mecanismo de acción y su relación con otras proteínas implicadas en el proceso de degradación de peptidoglicano.

■ **CbpL:**

- Cristalización y determinación estructural.
- Caracterización estructural del dominio de unión a colina y análisis comparativo con el resto de dominios de unión a colina previamente descritos.
- Caracterización funcional.



---

---

## MATERIAL Y MÉTODOS

---

---



---

# CRISTALOGRAFÍA DE RAYOS X

---

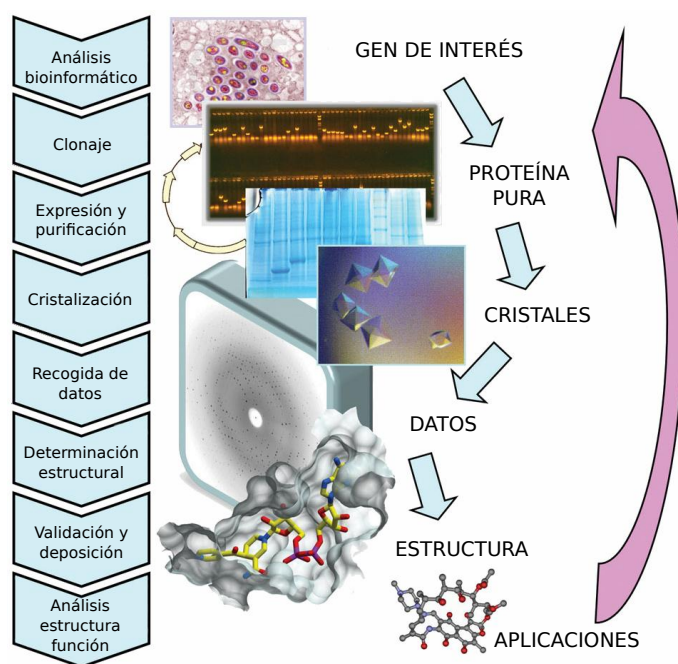
## Índice

---

3.1	Obtención de la proteína . . . . .	<b>26</b>
3.1.1	Clonación y expresión . . . . .	27
3.1.2	Purificación . . . . .	29
3.2	Cristalización de proteínas . . . . .	<b>31</b>
3.2.1	Nucleación y crecimiento cristalino . . . . .	31
3.2.2	Control del crecimiento cristalino . . . . .	32
3.2.3	Cristalización en presencia de ligandos . . . . .	33
3.2.4	Configuración de los experimentos de cristalización . . . . .	34
3.2.5	Búsqueda de condiciones iniciales y optimización . . . . .	36
3.2.6	Defectos en los cristales . . . . .	38
3.3	Difracción de rayos X y resolución estructural . . . . .	<b>39</b>
3.3.1	Configuración del experimento de difracción . . . . .	39
3.3.2	Obtención y análisis del patrón de difracción . . . . .	44
3.3.3	Determinación de la estructura tridimensional de la proteína . . . . .	49
3.3.4	Detección, análisis y resolución de datos maclados . . . . .	65

---

La técnica que se ha utilizado a lo largo de este trabajo para la resolución estructural de proteínas es la cristalografía de rayos X. Esta técnica permite, entre otras aplicaciones, determinar a escala atómica la estructura tanto de proteínas sencillas como de complejos macromoleculares o incluso virus. La técnica consta de varias etapas, resumidas en la Figura 3.1. En general, es necesario disponer de un cristal de proteína que, al interaccionar con un haz de rayos X, generará un patrón de difracción útil para determinar la estructura final de la proteína que conforma dicho cristal (Als-Nielsen & McMorrow, 2011).



**Figura 3.1:** Etapas básicas de la resolución estructural de proteínas mediante cristalografía de rayos X. Figura adaptada de Rupp, 2009.

A continuación se describen brevemente las diferentes etapas de la técnica, introduciendo los conceptos necesarios para la interpretación y discusión de los resultados obtenidos en este trabajo.

### 3.1. Obtención de la proteína

El primer paso de la resolución de estructuras mediante cristalografía de rayos X es obtener una muestra de proteína adecuada para su cristalización. Son requisitos indispensables que dicha muestra presente un alto grado de pureza (90 - 95 %), que sea estable, homogénea y que se pueda producir en cantidad suficiente para realizar tanto ensayos preliminares de cristalización como para obtener cristales finales de tamaño adecuado (Bergfors, 2009).

Para ello, se debe seguir un proceso de clonación, expresión y purificación que nos permita llegar del gen de interés a la proteína soluble. Cada proteína seguirá un proceso exclusivo de obtención en base a sus propiedades físico-químicas y biológicas, lo que supone que, en determinadas ocasiones, esta etapa actúe como cuello de botella. Además, ciertos comportamientos de la proteína durante los ensayos de cristalización se pueden deber a los



componentes utilizados durante el proceso de obtención, por lo que es habitual retroceder hasta esta fase tras comprobar que la cristalización o la resolución no ha sido efectiva.

### 3.1.1. Clonación y expresión

La etapa de clonación se inicia uniendo el DNA codificante para la proteína de interés a un vector, generalmente un plásmido, que a su vez se introducirá en un microorganismo adecuado para la expresión. Normalmente, el sistema de expresión va ligado a la utilización de una cepa específica de un microorganismo determinado, por lo que inicialmente se suelen realizar pruebas piloto de clonación y expresión en diferentes sistemas de uso general. En función de los resultados preliminares, se pueden modificar determinados elementos del sistema de expresión para favorecer la producción de la proteína de interés, como cambios en los sitios de restricción, en los marcadores de resistencia, en la cepa o incluso en el microorganismo.

Además, dependiendo de cada caso, puede interesar realizar una mutagénesis dirigida sobre determinados aminoácidos de la proteína. Las aplicaciones más habituales de esta técnica son inactivar una enzima para cristalizarla en presencia de su sustrato, modificar la solubilidad de regiones expuestas al solvente, evitar la flexibilidad de dominios o incluso crear puentes de unión entre diferentes proteínas que forman un complejo.

#### 3.1.1.1. Cepas bacterianas

La elección de la cepa bacteriana es fundamental para conseguir unos buenos resultados. Cada cepa tiene unas características únicas y su elección dependerá de varios factores tales como el propósito del experimento (etapa de clonación o de expresión), el vector escogido (tipo de sistema de inducción, marcador de selección compatible, etcétera) o el comportamiento de la proteína (de fácil expresión, tóxica para el microorganismo, insoluble, etcétera). Durante este trabajo se han utilizado *E. coli* DH5 $\alpha$  como cepa de clonación y *E. coli* BL21(DE3) como cepa de expresión.

La cepa DH5 $\alpha$  es una de las más utilizadas en la etapa de clonación ya que, debido a sus propiedades, se incrementa la eficiencia de transformación y la estabilidad del plásmido. La mutación *endA1* inactiva una endonucleasa intracelular que degrada el DNA plasmídico. RecA impide la recombinación homóloga, reduciendo así la posibilidad de multimerización del plásmido y de deleciones. La mutación *hdsR17* elimina la endonucleasa de restricción del sistema de restricción/modificación *EcoKI*, evitando que se degrade el DNA carente de la metilación de *EcoKI*. La deleción de los aminoácidos 11 a 41 de la  $\beta$ -galactosidasa (*lacZ*) permite que se pueda realizar una selección clonal basada en color. Otras mutaciones aportan a la cepa una mayor resistencia al peróxido de hidrógeno y al ácido nalidíxico.

La cepa BL21(DE3), utilizada en la etapa de expresión, contiene el fago lisógeno  $\lambda$  (DE3), que incluye el gen *lacI* y el gen para la RNA polimerasa del fago T7 bajo el control del promotor/operador *lacUV5*. En ausencia de IPTG (isopropil- $\beta$ -D-1-tiogalactopiranosido), el producto del gen *lacI* se une al operador de *lacUV5*, lo que impide la transcripción de la RNA polimerasa. Cuando se induce la expresión con IPTG, este compuesto se une al represor *lacI* impidiendo que se una al operador y permitiendo así la expresión de la RNA polimerasa y de

la proteína recombinante codificada en el vector de expresión bajo el control del promotor de T7. Además, las cepas derivadas de *E. coli* B, como BL21(DE3), carecen de la proteasa *lon* y de la proteasa *ompT* de la membrana externa, que podrían degradar las proteínas durante su expresión. Por último, la mutación *hsdS* permite una transformación eficiente de DNA sin metilar y la mutación *gal* impide la utilización de galactosa por parte de la célula.

### 3.1.1.2. Vectores de expresión

La elección del vector es clave para la obtención de una cantidad adecuada de proteína. A la hora de elegir un vector hay que tener en cuenta varios factores, como la longitud del DNA que se va a insertar, el número de copias del vector, la compatibilidad con otros vectores, el marcador de selección, los sitios de restricción presentes y la compatibilidad de éstos con el inserto, el número de sitios de clonación presentes en caso de querer expresar varias proteínas, la presencia o ausencia de diversas etiquetas para facilitar la purificación de la proteína una vez expresada y el sistema de expresión utilizado.

Durante este trabajo se ha utilizado el vector pET-28b(+) (Novagen, Alemania). Los vectores del sistema pET contienen el origen de replicación pBR322, que produce de 20 a 60 copias del plásmido por cada célula. Dichos plásmidos proporcionan un control estricto de la expresión de la proteína. Para ello contienen el promotor de T7 seguido del operón *lac*. También contienen el gen *lacI*, represor controlado por su propio operador. De este modo, el *lacI* actuará simultáneamente sobre la célula, impidiendo que se produzca RNA polimerasa de T7, y sobre el gen de interés, impidiendo que se exprese a pesar de la presencia de RNA polimerasa de T7, a no ser que se añada IPTG y el represor *lacI* se libere del operón *lac*, permitiendo así que el gen se transcriba.

El plásmido pET-28b(+) contiene un único sitio de clonación múltiple bajo el control del promotor de T7. Presenta la posibilidad de añadir etiquetas de hexahistidina o T7 en el extremo N-terminal de la proteína expresada y/o etiqueta de hexahistidina en su extremo C-terminal para facilitar su purificación. El plásmido puede seleccionarse mediante el gen de resistencia a kanamicina que contiene.

### 3.1.1.3. Medios de cultivo

Existe un gran número de medios de cultivo que cumplen diferentes funciones (selección o crecimiento de distintos tipos de bacterias) y cuya elección depende en gran medida del organismo que se vaya a crecer. Muchas bacterias son auxótrofas para algún aminoácido o tienen requerimientos nutricionales específicos. Los medios pueden tener una composición definida y conocida o ser complejos, en caso de que se preparen a partir de extractos de materiales animales o vegetales digeridos. No se conoce la composición específica de los medios complejos, pero contienen todos los nutrientes necesarios para el crecimiento de la bacteria. El medio escogido es muy importante para mantener a las bacterias en perfectas condiciones y conseguir un rendimiento de expresión proteica adecuada. Pueden añadirse aditivos como sacarosa, glucosa, etanol, grupos prostéticos o cofactores al medio para aumentar la solubilidad de la proteína.

A lo largo de este trabajo se utilizó medio Luria-Bertani (LB) para el cultivo líquido de bacterias. Dicho medio contiene 10 g/L de triptona, 5 g/L de extracto de levadura y 10 g/L

de NaCl. Para los cultivos sólidos se usó LB-agar, con la misma composición que el LB líquido pero suplementado con 15 g/L de agar bacteriológico. Para la obtención de células competentes se usó medio TSS (Chung *et al.*, 1989), compuesto de LB con 12 % (m/v) PEG 8000, 5 % (v/v) DMSO, 25 mM NaCl y 1 % (v/v) glicerol. El NaCl es fundamental para proporcionar a las células competencia química, mientras que el DMSO incrementa la competencia y el PEG y el glicerol ayudan a crioproteger las células cuando se congelan a -80 °C.

#### **3.1.1.4. Expresión de proteínas**

Para expresar la proteína se crece un cultivo de células transformadas con el plásmido que contiene el gen de interés en un medio de cultivo adecuado, con un antibiótico determinado, durante varias horas y a una temperatura óptima de crecimiento y con agitación. Al alcanzar una densidad óptica adecuada (generalmente durante la fase exponencial de crecimiento) se induce la expresión de la proteína mediante la adición de algún compuesto al medio, continuando el crecimiento del cultivo durante algunas horas más. En caso de algunas proteínas con problemas para el plegamiento es necesario enfriar el cultivo en hielo antes de inducir la expresión y, posteriormente, dejar crecer las células a unos 16 °C en agitación durante un periodo más largo, para favorecer una expresión más lenta y un mejor plegamiento. Otro método para mejorar la cantidad de proteína soluble es reducir la cantidad de inductor para conseguir una expresión más lenta y dar más tiempo al plegamiento. Una vez finalizada la etapa de crecimiento y expresión, el cultivo se centrifuga, se resuspende en un tampón de lisis y se congela para su conservación hasta la etapa de purificación.

#### **3.1.2. Purificación**

Una vez obtenido un cultivo que ha expresado la proteína de interés se inicia un proceso de purificación en el que se extrae la proteína de la bacteria, con el mayor grado de pureza posible. Para ello se dispone de varias técnicas. Es necesario tener en cuenta que en cada paso de la purificación, se obtiene proteína cada vez más pura, pero se pierde gran cantidad de la proteína total. Así pues, el rendimiento disminuirá si se añaden demasiados pasos a la purificación. Por ello es conveniente llevar un control de la calidad y cantidad de la muestra resultante tras cada paso de purificación.

##### **3.1.2.1. Lisis celular**

En primer lugar, se deben romper las células que han expresado la proteína de interés. La elección del tampón de lisis influye en gran medida en el rendimiento de proteína soluble obtenida. Un buen tampón de lisis debe tener un pH compatible con la proteína que se vaya a purificar y, a ser posible, compatible con el primer paso de cromatografía que se vaya a realizar. Los tampones de lisis suelen usarse en concentraciones que van de 20 a 50 mM. Además, se suelen añadir otros compuestos al tampón, como sales a concentraciones de 50 a 150 mM para mantener la fuerza iónica del medio y aumentar así la solubilidad y estabilidad de la proteína. Además suele llevar un 5 - 15 % de glicerol para estabilizar la proteína y evitar interacciones inespecíficas con otros compuestos celulares. Dependiendo de la proteína pueden añadirse quelantes, inhibidores de proteasas, ligandos o detergentes y agentes reductores para conseguir un mejor resultado.

El método de lisis elegido tampoco es arbitrario. Para cristalografía de proteínas es preferible evitar el uso de lisozima para realizar la lisis, debido a su gran facilidad para cristalizar. Uno de los métodos más utilizados es la sonicación, aunque hay que tener en cuenta la liberación de grandes cantidades de calor que podrían dañar la proteína. Por ello, es necesario conservar la muestra en hielo e intercalar ciclos cortos de sonicación con ciclos de incubación en hielo. Para grandes volúmenes son más adecuadas técnicas como la prensa de French, en la que las células pasan por una cámara muy estrecha empujadas por un émbolo a alta presión.

Una vez se ha producido la lisis, se realiza una centrifugación para separar las proteínas del resto de compuestos celulares. Se analizan tanto el precipitado como el sobrenadante por medio de una electroforesis desnaturalizante en gel de poliacrilamida (SDS-PAGE) llevada a cabo según el método de Laemmli (Laemmli, 1970). En este trabajo se utilizaron geles de poliacrilamida a una concentración del 12 % (m/v) para el separador y del 4 % (m/v) para el concentrador.

### **3.1.2.2. Cromatografía de afinidad**

Esta técnica se basa en la afinidad de una proteína concreta por un compuesto unido a una resina, como puede ser un anticuerpo, un ligando o, en el caso de las proteínas con cola de poli-His, Ni-NTA. La interacción reversible entre dicho compuesto y la proteína permite la separación de la proteína del resto de compuestos del extracto proteico. La principal ventaja de la técnica es la gran selectividad que ofrece. Para eluir la proteína de interés debe revertirse la interacción con la columna. Dicha elución puede ser específica (añadiendo un ligando que compita con la resina o con la proteína) o inespecífica (aumentando la fuerza iónica o cambiando el pH).

Una de las cromatografías de afinidad más ampliamente usadas es la resina de Ni-NTA. Dicha resina tiene una afinidad muy alta por las colas de histidina que se añaden a múltiples proteínas durante su clonación. El níquel interacciona con el grupo imidazol de las histidinas y retiene a la proteína en la matriz. Para la elución se añaden concentraciones de imidazol que van desde 50 mM hasta 1 M, compitiendo con la resina.

### **3.1.2.3. Cromatografía de intercambio iónico**

Las cromatografías de intercambio iónico separan las proteínas basándose en las diferencias de carga en la superficie de las distintas proteínas, produciéndose una interacción entre grupos polares cargados de la superficie de la proteína y de la resina de la columna. Para que una proteína esté lo suficientemente cargada para interaccionar con la columna debe estar disuelta en un tampón que presente un pH al menos dos puntos por encima o por debajo del pI de la proteína, en función de la carga que se quiera conseguir. Las proteínas con cargas superficiales opuestas a las de la resina empleada son retenidas, mientras que el resto son eluidas. Las proteínas retenidas son posteriormente eluidas gradualmente usando un gradiente creciente de NaCl en el mismo tampón.

Existen dos tipos principales de columnas de intercambio iónico, a saber, de intercambio aniónico (Q), que retienen proteínas cargadas negativamente (pH mayor que el pI de la proteína),

y de intercambio catiónico (S), que retienen proteínas cargadas positivamente (pH menor que el pI de la proteína).

Como se verá en la Sección 6.2.2, en este trabajo se ha utilizado una columna de intercambio aniónico (Q Sepharose® High performance; Sigma-Aldrich) pero actuando como columna de afinidad.

#### **3.1.2.4. Cromatografía de exclusión molecular**

La cromatografía de exclusión molecular o filtración en gel separa moléculas en virtud de sus diferencias de tamaño. La capacidad separadora reside en el gel cuya matriz consta de un gran número de esferas porosas constituidas por largas cadenas de polímeros unidas entre sí, formando una extensa red por la que se desplaza la muestra.

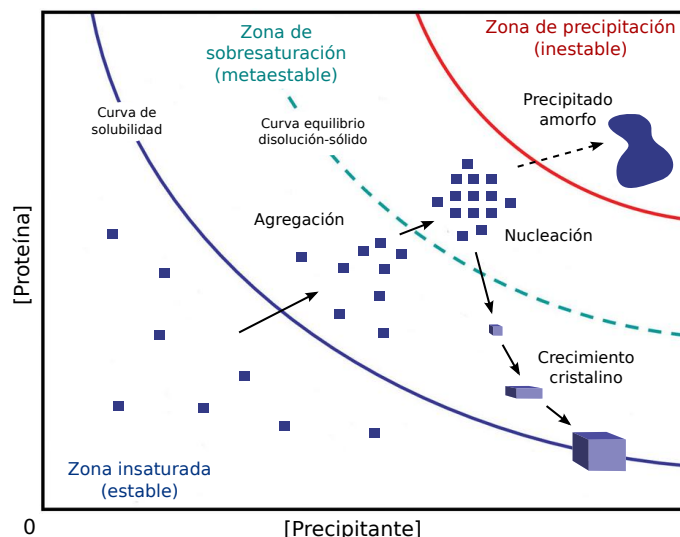
Cada gel se caracteriza por un rango de fraccionamiento que depende del tamaño de sus poros. Las columnas se describen mediante dos números que representan los pesos moleculares máximo y mínimo que pueden fluir por su interior. Las moléculas dentro de dicho rango presentan una mayor o menor facilidad para difundir al interior de las esferas en función de su tamaño, siendo extraídas fraccionadamente. Cuanto mayor es una molécula, menor es su capacidad para acceder al interior de las esferas del gel y, por tanto, menor es el volumen de líquido que se requiere para extraerla de la columna. Teniendo en cuenta que para las proteínas globulares el tamaño está, en general, relacionado con el peso molecular, este tipo de cromatografía separa las sustancias en función del peso molecular, obteniéndose primero las más pesadas. En el caso de proteínas fibrilares pueden existir desviaciones en los tiempos de elución, ya que la retención no estará directamente relacionada con su peso molecular.

### **3.2. Cristalización de proteínas**

Una vez obtenida la muestra de proteína, se debe proceder a su cristalización. Si se han cumplido las premisas de pureza, estabilidad y homogeneidad, la probabilidad de conseguir cristales se incrementa, mientras que una producción de peor calidad puede impedir la cristalización de la muestra. A continuación se describe el proceso de cristalización y se enumeran algunas de las técnicas más utilizadas actualmente para obtener cristales.

#### **3.2.1. Nucleación y crecimiento cristalino**

Tras la etapa de purificación, la proteína se encuentra disuelta en una solución que no ha alcanzado el punto de saturación, lo que significa que es termodinámicamente estable. Debido a ello, la proteína se mantendrá soluble siempre y cuando su concentración no alcance el punto de saturación. Cuando dicho punto se sobrepasa, la solución está sobresaturada, siendo termodinámicamente inestable. En este momento se inicia una precipitación que, si se controla, llevará a la obtención de cristales. En cambio, si la precipitación no se controla, se puede producir una precipitación masiva que de lugar a un precipitado amorfo, lejos del orden buscado en un cristal. Para estudiar el comportamiento de una solución que evoluciona hacia la sobresaturación se suelen utilizar los diagramas de solubilidad (Figura 3.2).



**Figura 3.2: Diagrama de fases del proceso de cristalización.** Las flechas indican el comportamiento habitual de la formación de cristales de proteína a partir de una muestra soluble. La zona metaestable se comporta como un gradiente en el que las proteínas se agregan progresivamente hasta formar núcleos.

En estos diagramas se enfrenta la concentración de proteína con una segunda variable del experimento, dejando las demás fijas. En las situaciones en las que se trabaja bajo la curva de solubilidad, la solución no ha alcanzado la saturación. El equilibrio se alcanza en la curva de solubilidad y, una vez sobrepasada, se inicia la sobresaturación del sistema, iniciándose la precipitación. Por tanto, el objetivo del experimento de cristalización es situar la proteína en solución en una región del diagrama de solubilidad que sobrepase la curva de solubilidad, pero lo suficientemente cercana a ella como para que la precipitación se produzca de forma ordenada y controlada. Si los niveles óptimos de sobresaturación son excedidos, la tasa de nucleación irá aumentando hasta hacerse masiva, generando desde dendritas hasta precipitados amorfos, en lugar de cristales individuales.

En cambio, el proceso de nucleación controlada provocará la agregación de moléculas de proteína en grupos individuales, denominados núcleos. Este hecho hace que disminuya la concentración de la proteína soluble, llevando el sistema a una situación en la que el crecimiento de los núcleos requiere menos energía que la formación de núcleos nuevos. Así pues, la nucleación se detiene y se inicia la etapa de crecimiento del cristal, durante la que la proteína soluble se une a la superficie de los núcleos preexistentes, aumentando su tamaño. Finalmente, el sistema se equilibra al llegar a la curva de solubilidad, deteniéndose el crecimiento del cristal.

### 3.2.2. Control del crecimiento cristalino

Ya que el proceso de cristalización puede descontrollarse y acabar en la formación de un precipitado amorfo, es fundamental manipular ciertas variables del experimento para obtener un ritmo de precipitación adecuado, permitiendo que las moléculas se ordenen en el espacio progresivamente, dando lugar a un cristal. A continuación se enumeran las principales variables de los experimentos de cristalización (Bergfors, 2009):

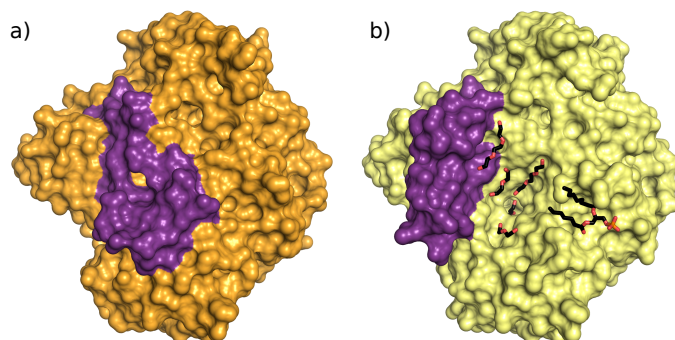
- **Concentración de proteína:** Generalmente se parte de una concentración base de 10 mg/mL, estadísticamente adecuada en la mayoría de los casos. No obstante, es conveniente realizar un test de precristalización (como la prueba PCT, *Pre-Crystallization Test*, Hampton Research) que indique si la proteína tiende a precipitar a una concentración cercana a la de partida.
- **Agentes precipitantes:** Se denominan agentes precipitantes cualquier tipo de molécula que pueda ayudar a la agregación progresiva de la proteína. Tanto su concentración como su naturaleza son críticas, aunque se ha probado su eficacia general en un amplio rango de casos. Lo más habitual es añadir una sal a baja concentración (0 a 300 mM) que aporte estabilidad y polímeros de diferente peso molecular, como el polietilenglicol (PEG). También suelen utilizarse alcoholes, que modifican la constante dieléctrica del solvente, o compuestos volátiles, cuya evaporación altera la solubilidad de la proteína paulatinamente.
- **pH:** La solubilidad de la proteína varía en función del pH al que se encuentre. Conviene ensayar un rango mínimo de pH de hasta entre 2 puntos por encima y por debajo del pI de la proteína. La concentración del tampón se suele situar sobre 100 mM.
- **Temperatura:** En general, las bajas temperaturas reducen la agitación térmica de la proteína y pueden favorecer la agregación y empaquetamiento. No obstante, en algunos casos conviene ensayar temperaturas elevadas, como en proteínas pertenecientes a organismos termófilos. Lo habitual es realizar experimentos a 18 °C y, en función del comportamiento, modificar la temperatura a 4 °C u otra adecuada.
- **Aditivos:** Se consideran aditivos las moléculas de bajo peso molecular y de baja concentración que, por su naturaleza, presentan capacidad de interacción con determinadas zonas de la proteína, pudiendo estabilizarlas y favorecer su empaquetamiento. Dependiendo de la proteína a cristalizar, se suelen utilizar azúcares, detergentes, aminoácidos, etcétera.

### 3.2.3. Cristalización en presencia de ligandos

En numerosas ocasiones es interesante obtener la estructura tridimensional de la proteína en presencia de algún ligando que interacciona con ella, como es el caso de sustratos y productos de reacción o de análogos no hidrolizables que permitan el análisis de estados de transición durante el mecanismo de reacción. Para obtener estas estructuras existen dos aproximaciones ampliamente utilizadas (Hassell *et al.*, 2007):

- **Cocrystalización:** La proteína se incuba en presencia del ligando antes de realizar los experimentos de cristalización. De este modo, el elemento a cristalizar es el complejo formado por la proteína y el sustrato.
- **Inmersión:** Una vez obtenidos los cristales de proteína, se sumergen en una disolución del ligando en la misma condición de cristalización, para evitar un desequilibrio osmótico que deteriore el cristal. Los canales de solvente presentes en los cristales de proteína permiten que el ligando circule por el interior del cristal y pueda acceder al sitio de unión.

Dependiendo de cada caso particular se utilizará diferente aproximación, variando los tiempos de incubación o inmersión y la concentración de ligando (Figura 3.3). En general, si el sitio de unión del ligando a la proteína está protegido del solvente, o si el ligando es de gran tamaño, una cocrystalización puede ser mejor aproximación que una inmersión, ya que el acceso del ligando a su sitio de unión podría encontrarse muy limitado por impedimentos estéricos. No obstante, en una cocrystalización pueden modificarse las condiciones de cristalización respecto a las condiciones de la proteína sin ligando, por lo que, en la práctica, conviene aprovechar ambos métodos para intentar obtener cristales del complejo proteína-ligando.



**Figura 3.3: Cristalización en presencia de ligandos.** La presencia de ligandos puede alterar la conformación de la proteína. En este ejemplo, la esterol esterasa de *Ophiostoma piceae* sufre un cambio conformacional notable al cristalizarse en presencia de ligandos. **a)** Conformación cerrada sin ligandos. **b)** Conformación abierta con ligandos (Gutiérrez-Fernández *et al.*, 2014).

### 3.2.4. Configuración de los experimentos de cristalización

Existen diversas técnicas que permiten llevar a cabo experimentos controlados de cristalización de macromoléculas. Aunque cada técnica presenta diversas variantes, generalmente se suelen agrupar en cuatro tipos (Chayen & Saridakis, 2008; Rupp, 2009):

- Difusión de vapor.
- Cristalización en lote (*batch*).
- Difusión sin barrera.
- Diálisis.

Pese a la existencia de esta clasificación, es habitual realizar configuraciones que combinan aspectos diferentes de dos o más técnicas. A continuación se describen las configuraciones más utilizadas durante el desarrollo de este trabajo.

#### 3.2.4.1. Difusión de vapor con aislamiento

En la cristalización por difusión de vapor con aislamiento, una gota que contiene la proteína en solución junto con un precipitante a determinada concentración es enfrentada a un volumen mayor de dicho precipitante a mayor concentración, todo ello aislado del exterior. Bajo estas condiciones, se producirá una evaporación de agua desde la gota hacia el precipitante. Así pues, el volumen de la gota irá disminuyendo, mientras que la proteína se va concentrando. Por tanto,

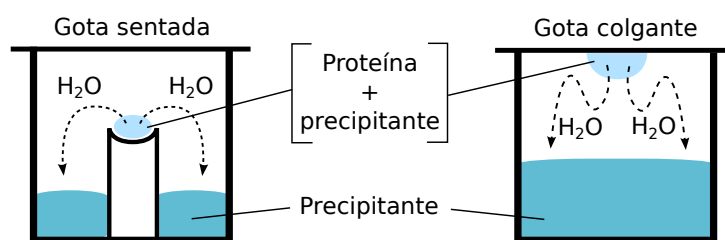


el sistema se acerca lentamente a la saturación y, posteriormente, a la sobresaturación, pasando por la zona de nucleación en la que se generarán los primeros núcleos. Finalmente tendrá lugar la etapa de crecimiento, en la que la proteína libre se unirá a los núcleos preformados dando lugar a cristales individuales.

En cierto momento, se alcanzará el equilibrio entre la gota y el precipitante, lo que conlleva la detención y finalización del experimento. Por este motivo, las condiciones de cristalización se ajustan en sucesivos experimentos, intentando acercarse lo máximo posible a la zona óptima de nucleación y crecimiento cristalino para la proteína y el precipitante utilizado.

Dentro de esta configuración experimental, se pueden utilizar dos montajes diferentes (Figura 3.4):

- **Montaje en gota sentada:** El precipitante a mayor concentración se sitúa en un pocillo que rodea una plataforma elevada sobre la que se sitúa la gota con la proteína y el precipitante diluido. El sistema se aísla del exterior mediante una película adhesiva impermeable que cubre todo el pocillo.
- **Montaje en gota colgante:** El precipitante a mayor concentración se sitúa en un pocillo sobre el que se coloca un cubreobjetos siliconado, en cuya cara interior transporta la gota con la proteína y el precipitante diluido. El aislamiento se consigue mediante la utilización de grasa de vacío entre el pocillo y el cubreobjetos.

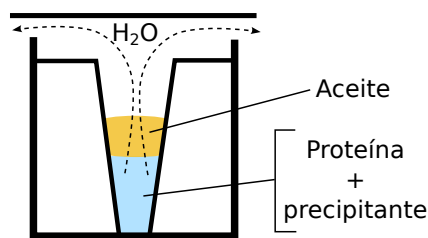


**Figura 3.4:** Cristalización por difusión de vapor con aislamiento. Configuración de gota sentada y de gota colgante.

#### 3.2.4.2. Difusión de vapor sin aislamiento

En las técnicas de cristalización por difusión de vapor sin aislamiento existe evaporación desde la gota de cristalización hacia el exterior sin que se alcance nunca el equilibrio, por lo que la evaporación es constante y la desecación inevitable (Figura 3.5). El sistema no se detiene, por lo que una vez formados los cristales se seguirá perdiendo agua, lo que provocará la precipitación del soluto, la pérdida de los cristales y la desecación de la gota. Así pues, estos experimentos deben ser constantemente vigilados para conocer el momento óptimo en el que los cristales alcanzan el tamaño máximo, antes de ser destruidos por la desecación.

Para evitar una desecación demasiado rápida, se utilizan aceites permeables basados en parafina y silicona que cubren la gota y reducen considerablemente la tasa de evaporación (Chayen, 1997). Este tipo de experimentos bajo aceite permeable suelen entrar en la clasificación de técnicas de cristalización en lote, si bien es cierto que se basa en la difusión de vapor. El motivo



**Figura 3.5:** Cristalización por difusión de vapor sin aislamiento.

es que, al no existir reservorio de precipitante, el material utilizado se reduce considerablemente y la puesta a punto del experimento es más rápida que en los métodos de difusión de vapor, lo que la hace ideal para ensayar condiciones iniciales en poco tiempo. Otra variante de la cristalización en lote utiliza aceites impermeables, lo que la sitúa fuera de las técnicas de difusión de vapor.

#### **3.2.4.3. Contradifusión en gel**

Es una variante del método de cristalización libre de barrera. En estas configuraciones, la proteína se pone en contacto directo con el precipitante sin que hayan sido mezclados previamente. La difusión de la proteína hacia el precipitante y viceversa conlleva la formación de un gradiente de concentración que permite explorar una zona amplia del diagrama de solubilidad.

En la práctica, la condición de cristalización debe ser conocida, por lo que esta configuración no es la más indicada para iniciar pruebas de cristalización sobre una proteína desconocida. Esta técnica es especialmente útil cuando una condición de cristalización ha devuelto cristales de baja calidad o precipitados cristalinos. En esos casos, las concentraciones de proteína y precipitante son las principales variables a explorar, lo que permite ajustar las tasas de nucleación y crecimiento cristalino para obtener cristales de mejor calidad (Ng *et al.*, 2003). También se está utilizando esta técnica en microfluídica, en las que la disponibilidad de proteína es más baja de lo habitual y se deben utilizar volúmenes pequeños (Leng & Salmon, 2009).

#### **3.2.5. Búsqueda de condiciones iniciales y optimización**

Un experimento de cristalización de una proteína se puede dividir, de forma genérica, en dos partes, a saber, búsqueda de condiciones iniciales y optimización de resultados (Cudney *et al.*, 1994; Benvenuti & Mangani, 2007).

En primer lugar, la proteína soluble se enfrenta a numerosas condiciones de cristalización diferentes, observando el tipo de precipitado que se genera en cada caso. Evaluando estos resultados se puede determinar qué precipitantes son más favorables, a qué concentración, pH, temperatura, etcétera (Luft *et al.*, 2011). Estas primeras pruebas se realizan con técnicas de alto rendimiento que permiten configurar numerosos experimentos en un tiempo relativamente corto. A lo largo de este trabajo se han realizado las búsquedas iniciales utilizando el robot de cristalización Innovadine Nanodrop<sup>TM</sup> ExtY y diversos paquetes comerciales de cristalización (Tabla 3.1), aplicando la técnica de difusión de vapor en aislamiento con configuración de gota sentada sobre placas Innovaplate<sup>TM</sup> SD-2. Por tratarse de ensayos preliminares, se utilizaron

volúmenes entre 250 – 500 nL de proteína y precipitante en gota, más 65  $\mu\text{L}$  de precipitante en el pocillo.

**Tabla 3.1:** Relación de paquetes comerciales de cristalización utilizados en diversos ensayos a lo largo de este trabajo.

NOMBRE	CASA COMERCIAL	REFERENCIA
Crystal Screen	Hampton Research	HR2-110
Crystal Screen 2	Hampton Research	HR2-112
INDEX	Hampton Research	HR2-144
Detergent Screen	Hampton Research	HR2-408
Additive Screen	Hampton Research	HR2-428
JCSG	Qiagen	130720
PACT Suite	Qiagen	130718
JBScreen Classic 1-4	Jena Biosciences	CS-101L/104L
JBScreen Classic 6	Jena Biosciences	CS-106L
Morpheus	Molecular Dimensions	MD1-46

Generalmente, los cristales que se obtienen en esta fase son pequeños y suelen presentar defectos e irregularidades. Por ello, tras la búsqueda de condiciones de cristalización iniciales, se realiza una etapa de optimización que, dependiendo de cada caso particular, puede conllevar más o menos pasos.

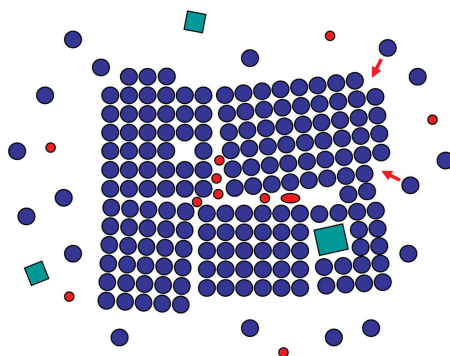
Una de las técnicas de optimización más habituales es el escalado, que consiste en aumentar los volúmenes de proteína y precipitante tanto en la gota (hasta 1 o 2  $\mu\text{L}$ ) como en el pocillo (entre 250 y 500  $\mu\text{L}$ ), con el objetivo de conseguir cristales de mayor tamaño que faciliten la toma de datos de difracción. Además, al mismo tiempo que se realiza el escalado se varían ligeramente las concentraciones de los componentes de la condición de cristalización, aumentando el rango de diferentes situaciones ensayadas. También es habitual realizar una variación de la relación de volúmenes entre la proteína y el precipitante en gota (1:2, 2:1, 1:3, etcétera). Con esto se busca alterar la concentración relativa de la proteína respecto al precipitante al mismo tiempo que se hace un escalado. Cambios en el pH y la temperatura, o la inclusión de aditivos o ligandos propios de la proteína, son variables a tener en cuenta en esta etapa. La optimización de los cristales obtenidos en este trabajo se ha realizado mediante técnicas de difusión de vapor en gota sentada, en gota colgante y técnicas de cristalización en lote bajo aceites permeables.

Por último, existe una técnica denominada *sembrado* que suele mejorar notablemente la calidad de los cristales (Bergfors, 2003). Si bien existen diferentes tipos de sembrado, la base de la técnica es común, consistente en obtener núcleos de un cristal preformado que sirvan como base para el crecimiento de nuevos cristales. Para ello, se tritura un cristal con defectos y se introduce en la solución una fibra natural, que arrastrará núcleos adheridos a ella. Esta fibra se sumergirá en una gota que contenga la proteína y la concentración de cristalización, liberando los núcleos provenientes del cristal anterior. De esta forma se puede realizar un proceso de crecimiento controlado sobre núcleos preformados. Una variante muy utilizada de este método consiste en aislar un cristal muy pequeño de la condición inicial y llevarlo a una gota con la

misma condición equilibrada y más proteína soluble, obteniendo un crecimiento en tamaño del cristal y no la formación de cristales nuevos.

### 3.2.6. Defectos en los cristales

El crecimiento de los cristales es un proceso delicado en el que es bastante frecuente que aparezcan defectos que alteren el orden y la periodicidad tan necesarios en cristalografía de rayos X (Figura 3.6). Entre los defectos más comunes están las inclusiones y las fracturas, que generan dominios ligeramente desviados entre sí. Estas desviaciones provocan lo que se denomina *mosaicidad*, un parámetro a tener en cuenta durante el procesamiento de los datos de difracción ya que altera la forma y tamaño de cada reflexión (Bellamy *et al.*, 2000).



**Figura 3.6: Defectos en el crecimiento de los cristales.** La aparición de impurezas en el experimento de cristalización da lugar a inclusiones y fracturas en el cristal. Si la mosaicidad es muy elevada, la proteína no podrá ser resuelta (Rupp, 2009).

#### 3.2.6.1. Cristales maclados

Como se ha dicho, las impurezas pueden alterar el empaquetamiento de un cristal y elevar la mosaicidad hasta límites irresolubles, pero un problema mucho más grave es la aparición de maclas.

Las maclas surgen cuando, bien por fracturas, bien por diferentes empaquetamientos, se generan dos o más dominios macroscópicos con diferente orientación dentro de un mismo cristal. El hecho de que el defecto sea macroscópico no se traduce en un aumento de la mosaicidad sino en la aparición de dos o más patrones de difracción superpuestos, cada uno correspondiente a una orientación de las que componen la totalidad del cristal. En la práctica, equivale a tener un cristal formado por dos o más cristales de la misma proteína con diferente orientación (Rupp, 2009).

En ocasiones existe una frontera física clara entre las diferentes orientaciones presentes en un cristal, por lo que mediante esfuerzo mecánico se puede inducir su rotura, separando los diferentes dominios que se comportarán como cristales únicos. Este tipo de macla, denominada *de contacto*, se contrapone a las denominadas maclas *interpenetradas*, en las que las diferentes orientaciones se encuentran entrelazadas a lo largo de todo el cristal, sin que sea posible su separación (Giacovazzo *et al.*, 2011). Como se verá más adelante (Sección 5.3), la determinación

estructural de una proteína cuyos cristales están maclados se puede complicar hasta el punto de que sea imposible su resolución.

En la Sección 3.3.4 se analizan los diferentes tipos de macla y se describen los aspectos a tener en cuenta para detectar esta patología durante el procesamiento de datos, así como algunos aspectos clave a la hora de intentar su resolución.

### 3.3. Difracción de rayos X y resolución estructural

Una vez que se ha conseguido un cristal formado por la proteína, buscamos conocer su estructura tridimensional. Para ello debemos recurrir al fenómeno de difracción que se produce al incidir un haz de rayos X sobre el cristal, generando un patrón de haces difractados hacia diferentes direcciones y con diferentes intensidades (Figura 3.7). Entender el fenómeno de la difracción y su relación con la simetría interna del cristal es fundamental para llegar a obtener el modelo atómico de la molécula de interés. Así pues, en las siguientes páginas se realiza un breve análisis tanto del aspecto técnico de la configuración del experimento como del proceso de difracción de la luz llevado a cabo tras la interacción con el cristal y su procesamiento para obtener la estructura final.



**Figura 3.7: Esquema general del experimento de difracción.** El experimento de difracción de rayos X se divide en 5 etapas fundamentales bien determinadas: 1) Generación y emisión de los rayos X. 2) Interacción con el cristal y difracción. 3) Detección. 4) Análisis y procesamiento de los datos. 5) Obtención de la estructura. Como se verá más adelante, el procesamiento de datos y la resolución se subdivide a su vez en diferentes etapas que varían dependiendo de cada caso.

#### 3.3.1. Configuración del experimento de difracción

Para poder llevar a cabo un experimento de difracción de rayos X sobre un cristal de proteína se requieren los siguientes elementos:

- **Fuente de rayos X:** Elemento encargado de generar los rayos X que van a servir para irradiar el cristal de proteína.
- **Soporte del cristal:** Elementos de sujeción y protección del cristal.
- **Detector:** Elemento de detección de la radiación difractada por el cristal.

A continuación se detallan los aspectos más importantes de cada uno de estos elementos.

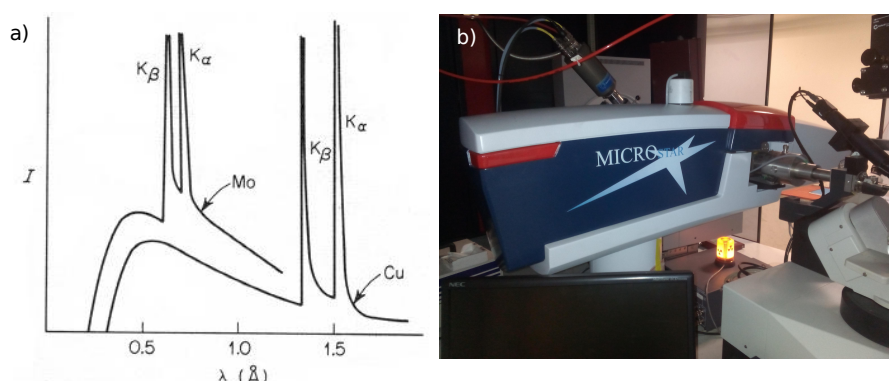
##### 3.3.1.1. Fuentes de rayos X

Los rayos X son radiaciones electromagnéticas cuya longitud de onda se sitúa en el rango de  $0,1 - 100 \text{ \AA}$  (energías de  $123,98 \text{ keV}$  a  $0,12398 \text{ keV}$ ). En los experimentos de cristalografía de

rayos X para la resolución de estructuras de macromoléculas se utiliza una banda que abarca, aproximadamente, desde los 0.6 Å ( $\simeq$  20 keV) hasta los 2,5 Å ( $\simeq$  5 keV). La elección de este rango se debe a que equivale a la magnitud de los enlaces existentes entre los átomos de las moléculas que conforman el cristal, por lo que la interacción de esta radiación con la materia es capaz de devolver información a resolución atómica (Rupp, 2009; Als-Nielsen & McMorrow, 2011).

La obtención de rayos X con estas características puede llevarse a cabo tanto en generadores de ánodo rotatorio de laboratorio como en instalaciones de radiación sincrotrón:

- **Generador de ánodo rotatorio:** Se establece una diferencia de potencial de alto voltaje entre un filamento incandescente cargado negativamente (cátodo) y un metal puro, generalmente molibdeno o cobre, cargado positivamente (ánodo). Los electrones que viajan desde el filamento hasta el ánodo provocan una reorganización en los orbitales atómicos del metal, liberando fotones de energías dentro del rango útil para la cristalografía de proteínas (Figura 3.8a). El hecho de que el ánodo sea un cilindro en continua rotación magnifica la superficie de contacto de los electrones que viajan desde el filamento, permitiendo un aumento en la potencia de los rayos X generados. El generador de ánodo rotatorio utilizado durante este trabajo fue un MicroStar de 2,7 kW con microfoco de 100  $\mu\text{m}$  (Bruker) (Figura 3.8b).

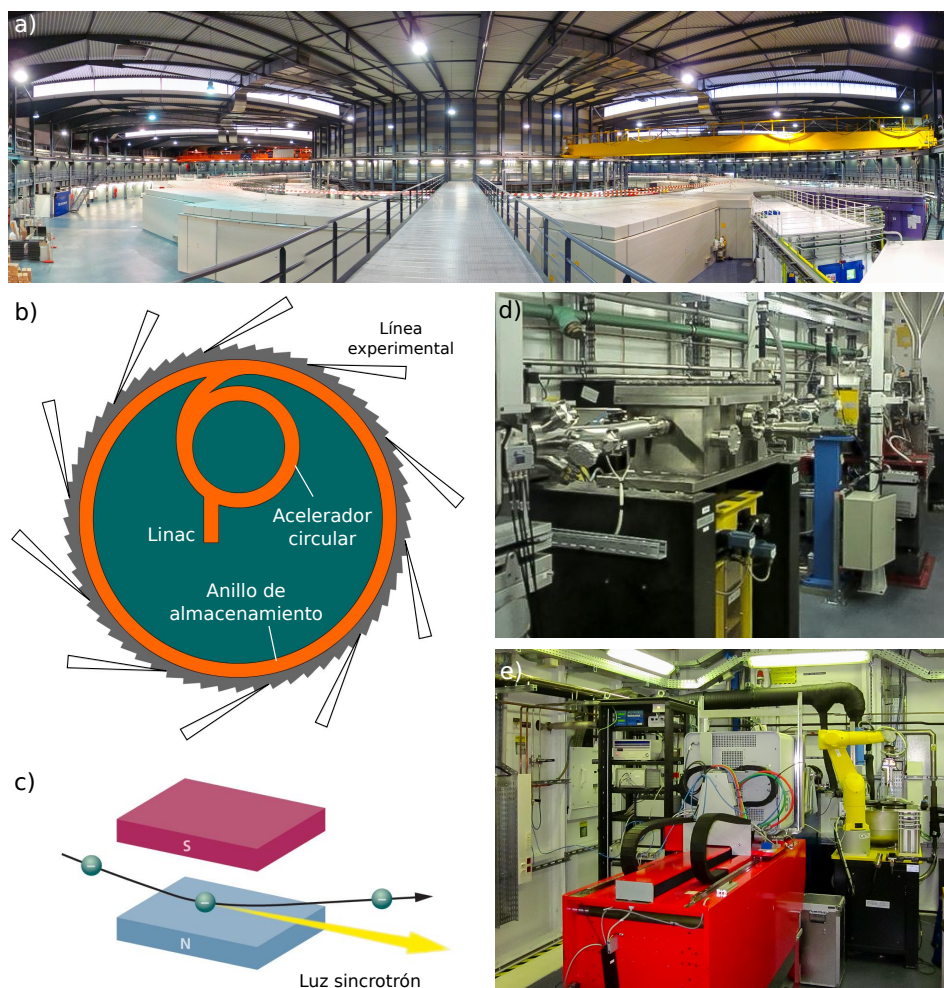


**Figura 3.8: Espectros de emisión de cobre y molibdeno y generador de ánodo rotatorio.** a) Superposición de los espectros de emisión de cobre y molibdeno. Los picos se han truncado para facilitar la representación. No obstante, la emisión de la banda  $K\alpha$  es notablemente más intensa que la emisión  $K\beta$ , motivo por el que se prefiere el uso de la  $K\alpha$ . b) Generador de ánodo rotatorio MicroStar de 2,7 kW, perteneciente al Instituto de Química-Física Rocasolano (IQFR-CSIC).

- **Radiación sincrotrón:** Una instalación de radiación sincrotrón integra un acelerador de electrones lineal en el que se inyectan paquetes de electrones que sufren una aceleración inicial, un anillo de aceleración que eleva la energía de dichos paquetes hasta valores cercanos a 6 GeV, un anillo de almacenamiento de varios metros de diámetro por el que viajan los paquetes de electrones de alta energía y una serie de líneas experimentales en posición tangencial al último anillo (Figuras 3.9a y 3.9b). Los paquetes de electrones tienden a viajar en línea recta y, si son desviados, emiten un gran espectro de radiación electromagnética de alto brillo (Figura 3.9c). Por este motivo, las líneas experimentales se



sitúan de manera tangencial al anillo de almacenamiento, recogiendo la radiación liberada por los electrones en el giro. Dicha radiación se filtra en la denominada cabaña óptica de la línea experimental (Figura 3.9d) con la finalidad de obtener las energías necesarias para el experimento de difracción. Por último, la radiación llega a la cabaña experimental, donde alcanza el cristal y el detector (Figura 3.9e). Las diferentes líneas de radiación sincrotrón utilizadas a lo largo de este trabajo aparecen enumeradas en la Tabla 3.2.



**Figura 3.9: Instalación de radiación sincrotrón.** a) Vista panorámica del sincrotrón ALBA (Cerdanyola del Vallès, España). b) Esquema de las partes fundamentales de un sincrotrón y su disposición espacial. c) Esquema simplificado del funcionamiento de un dispositivo de desviación de paquetes de electrones basado en imanes. La radiación sincrotrón se libera continuando la trayectoria inicial de los electrones. d) Cabaña óptica de la línea BL13 (Xaloc), perteneciente al sincrotrón ALBA. e) Cabaña experimental de la línea BL13 (Xaloc), mostrando el detector (Pilatus 6M) y el sistema de montaje de cristales robotizado (CATS).

### 3.3.1.2. Soporte y protección del cristal

Para realizar una medida estable, los cristales deben estar ubicados en un soporte que los mantenga en una posición fija y los permita rotar. Además, deben estar protegidos de la propia radiación X, ya que no sólo generan el patrón de difracción sino que también provocan numerosos daños que se van acumulando a lo largo de la medida y deterioran el resultado final (Garman & Owen, 2006; Ravelli & Garman, 2006; Garman, 2010).

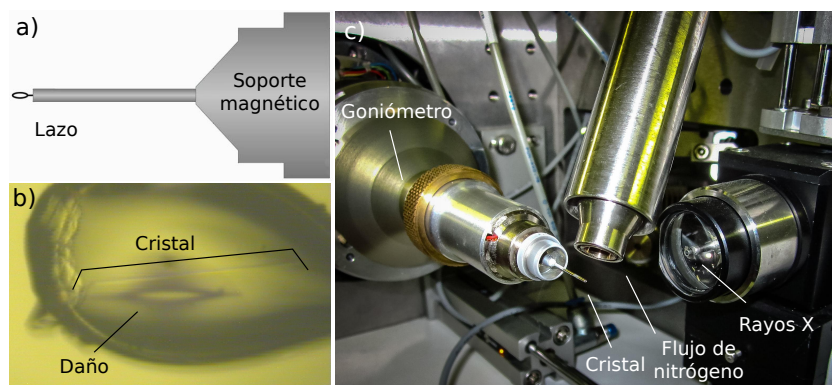
**Tabla 3.2:** Relación de líneas de radiación sincrotrón utilizadas durante este trabajo.

SINCROTRÓN*	LÍNEA	DETECTOR	LOCALIZACIÓN
SLS	X06SA (PXI)	MARmosaic 225**	Villigen, Suiza
SLS	X06SA (PXI)	Pilatus 6M	Villigen, Suiza
SLS	X06DA (PXIII)	Pilatus 2M	Villigen, Suiza
ESRF	ID14-1	ADSC Q210 CCD	Grenoble, Francia
ESRF	ID14-4	ADSC Q315r CCD	Grenoble, Francia
ESRF	ID23-1	ADSC Q315r CCD	Grenoble, Francia
ESRF	ID29	Pilatus 6M	Grenoble, Francia
DLS	I03	Pilatus 6M	Oxfordshire, Reino Unido
ALBA	BL13 (Xaloc)	Pilatus 6M	Cerdenyola del Vallès, España

\* SLS: *Swiss Light Source*. DLS: *Diamond Light Source*. ESRF: *European Synchrotron Radiation Facility*.

\*\* Se utilizó el microdifractómetro de MAATEL MkII (MD2).

Los cristales de proteína se recogen de la placa de cristalización con un lazo de *nylon* de tamaño similar al de cada cristal en particular (habitualmente de  $0,025\ \mu\text{m}$  a  $0,5\ \mu\text{m}$ ; Figura 3.10a). Para reducir el daño que sufre el cristal por la radiación incidente (Figura 3.10b), el cristal se sumerge en una gota que contiene solución crioprotectora antes de someterle a una corriente de nitrógeno cuya temperatura suele situarse en 100 K. Finalmente, el lazo se monta sobre un goniómetro que permitirá su centraje y rotación durante la recogida de los datos (Figura 3.10c).



**Figura 3.10:** Montaje de un cristal en un difractómetro de rayos X o línea de radiación sincrotrón. a) Esquema de los lazos de *nylon* utilizados como soporte para los cristales. b) Cristal de la proteína DacB, resuelta en este trabajo, atravesado por el haz de rayos X tras tomar una imagen de 0,2 segundos de exposición, sin crioprotger. c) Montaje general del cristal en el goniómetro de la cabaña experimental de BL13 (Xaloc, ALBA), centrado respecto a la fuente de rayos X y protegido por la corriente de nitrógeno.

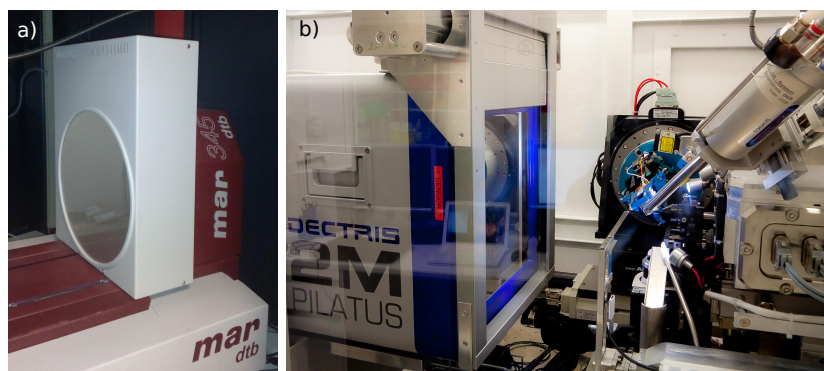
En ocasiones es interesante realizar alguna medida de referencia a temperatura ambiente para saber cómo se comporta el cristal sin crioprotger y poder extrapolar los posibles daños por radiación a parámetros como la temperatura del nitrógeno o el tipo y concentración de crioprotector utilizado. Los crioprotectores más comunes son glicerol, MPD y Paratone<sup>®</sup>, aunque en algunos casos se pueden utilizar criosales o componentes propios de la condición de cristalización que se comportan como crioprotectores, como algunos tipos de PEG (Garman & Mitchell, 1996; Robinson *et al.*, 2000; Berejnov *et al.*, 2006).



### 3.3.1.3. Detección de rayos X

Finalmente, los rayos X difractados se deben registrar, para lo que se utiliza un detector. Inicialmente se utilizaba la clásica película fotográfica, pero esa tecnología fue rápidamente desbancada por los detectores de área, que supusieron una revolución en la captura de imágenes de difracción. Existen tres tipos de detectores de área que son utilizados actualmente en este tipo de experimentos:

- **Detectores de fosforescencia:** Más conocidos como *imaging plates*, consisten en una superficie fotosensible formada por cristales de diferentes compuestos que se excitan con la llegada de los fotones, manteniendo el estado de excitación durante un tiempo. Posteriormente, un láser barre la superficie excitada, registrando la información, y, finalmente, se procede a un borrado del estado de la superficie para proceder a otra lectura. Inicialmente, los tiempos de lectura eran muy lentos y la sensibilidad baja, pero actualmente, los detectores de este tipo han sufrido una gran evolución que les permite seguir siendo utilizados en numerosos laboratorios. No obstante, debido a los aún bajos ritmos de lectura y a la necesidad de borrar la información periódicamente, su uso en sincrotrones es prácticamente nulo, por lo que su uso mayoritario se encuentra asociado a generadores de ánodo rotatorio. (Figura 3.11a). Durante este trabajo, el detector de fosforescencia que se ha utilizado para las medidas provenientes del ánodo rotatorio ha sido un Mar345dtb (MarResearch).
- **Detectores CCD y CMOS:** En lugar de almacenar la información temporalmente para una lectura posterior, como sucede con los *imaging plates*, los detectores CCD (*Charged-Coupled Device*) y CMOS (*Complementary Metal-Oxide-Semiconductor*) utilizan numerosas celdillas fotoeléctricas que generan un impulso eléctrico proporcional al número de fotones que llegan a su superficie. Este sistema es más sensible a la radiación que los detectores de fosforescencia, por lo que acorta los tiempos de exposición de minutos a segundos y permite registrar reflexiones más débiles.
- **Detectores HPAD:** Los últimos modelos de detectores son los HPAD (*Hybrid Pixel Array Detectors*), que mejoran notablemente la calidad de los detectores CCD/CMOS. La diferencia fundamental es que, mientras que un detector CCD/CMOS convierte los rayos X en luz visible detectable, la tecnología HPAD detecta directamente los fotones de los rayos X mediante píxeles de silicio, sin necesidad de realizar una conversión a otra energía. Carecen prácticamente de ruido de fondo y son capaces de discriminar cada fotón de cada reflexión, acortando aún más los tiempos de exposición, que se sitúan en fracciones de segundo (Figura 3.11b). Además, la función de dispersión de punto es significativamente más precisa que en otros detectores, por lo que la topología de cada reflexión se determina con mayor precisión. Los detectores utilizados durante este trabajo se muestran en la Tabla 3.2.



**Figura 3.11: Detectores utilizados en cristalografía de rayos X.** a) Detector Mar345dtb de tipo *imaging plate*, perteneciente al Instituto de Química-Física Rocasolano (IQFR-CSIC). b) Detector Pilatus 2M (Dectris) de tipo HPAD, perteneciente a la línea X06DA (PXIII) del SLS.

### 3.3.2. Obtención y análisis del patrón de difracción

La base de la cristalografía de rayos X es utilizar un cristal de proteína como una red de difracción, lo que permite obtener un patrón de difracción de rayos X único para cada cristal (y, por ende, único para cada proteína). Ya que cada proteína se empaqueta en un cristal único que a su vez genera un patrón único, es evidente que a cada patrón de difracción le corresponde una única estructura proteica. Así pues, interesa conocer el proceso de formación de dicho patrón para, posteriormente, siguiendo el camino inverso, determinar la estructura de la proteína que lo ha generado.

Además, la aparición progresiva de diversos programas informáticos para el procesamiento de los patrones de difracción y la obtención de la estructura final ha facilitado enormemente esta tarea, aumentando tanto la velocidad de resolución de estructuras como su número total. Recientemente se superó el valor de 100000 estructuras depositadas en el PDB (*Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University; Berman et al., 2000*) y su número sigue creciendo diariamente.

A continuación se describen brevemente algunos de los aspectos teóricos básicos del proceso de difracción de rayos X por un cristal de proteína, fundamentales para comprender cómo se obtiene un patrón de difracción y qué información se puede extraer de él.

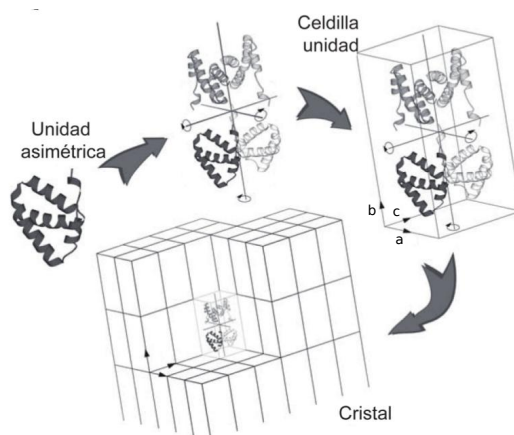
#### 3.3.2.1. Dispersión de la radiación y factor de dispersión atómico

La difracción se puede definir como la interferencia que se produce al dispersarse la luz a causa de una serie de dispersores ordenados en el espacio. Esto hace que, en ocasiones, se hable de la difracción como una dispersión cooperativa. En el caso de las proteínas, los dispersores son los átomos de las proteínas que conforman el cristal. Cada átomo dispersa la luz en función de su factor de dispersión atómico,  $f$ , valor que se corresponde con la amplitud de la onda dispersada tras la interacción con dicho átomo (Rupp, 2009; Giacovazzo *et al.*, 2011). Como se verá más adelante (Sección 3.3.3.1), este parámetro es de crucial importancia para el cálculo de la densidad electrónica de la proteína, que nos llevará a la determinación de la estructura final.

El hecho de que la dispersión de los rayos X por un cristal genere un patrón de difracción se debe a que los átomos, que actúan como dispersores, se encuentran ordenados de forma periódica en el espacio, generando una red de interferencia tridimensional. Así pues, es fundamental encontrar un sistema que permita describir la simetría que existe entre los elementos que forman un cristal.

### 3.3.2.2. Simetría, red directa y red recíproca

Debido a que las proteínas de un cristal se encuentran ordenadas de forma periódica, es posible determinar una zona mínima del cristal que se repite por traslación en las tres direcciones espaciales. A esta zona se le denomina *celdilla unidad*, la cual suele contener varias moléculas de proteína a su vez relacionadas entre sí. La celdilla unidad es un paralelepípedo caracterizado por la longitud de sus tres aristas ( $a$ ,  $b$  y  $c$ ) y sus tres ángulos ( $\alpha$ ,  $\beta$  y  $\gamma$ ). En el interior de dicho paralelepípedo se sitúan las proteínas. Las relaciones entre las distintas moléculas de proteína que hay en cada celdilla unidad surgen de la aplicación de elementos de simetría como ejes de rotación, planos de reflexión y planos de deslizamiento. Así pues, debido a la aplicación de dichas operaciones de simetría, se define una *unidad asimétrica*, que es el volumen mínimo cuya determinación es necesaria para poder extrapolar la información al resto de la celdilla unidad y, con ello, a todo el cristal (Figura 3.12).

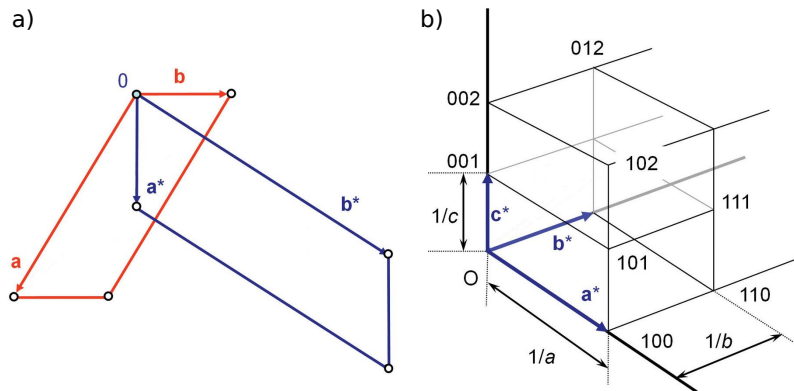


**Figura 3.12: Unidad asimétrica, celdilla unidad y cristal.** Representación esquemática de los elementos de organización de las proteínas dentro de un cristal. Figura adaptada de McPherson, 2009.

Las diferentes operaciones de simetría que se pueden aplicar a una molécula para formar un cristal dan lugar a 230 grupos espaciales diferentes, es decir, 230 formas distintas de repetir esa molécula en el espacio. En el caso de las proteínas, el número de grupos espaciales se reduce a 65 debido a que, por la quiralidad que presentan estas moléculas, se restringen ciertas operaciones de simetría.

La posición de cada átomo de la proteína dentro de la celdilla unidad se determina mediante un vector de posición individual. Esto permite analizar el cristal como si de un espacio vectorial se tratara. Dicho espacio vectorial recibe el nombre de espacio real y, la red obtenida en él, *red directa*. A partir de la red directa se puede obtener una *red recíproca*, enmarcada en un espacio vectorial recíproco, cuyos elementos están directamente relacionados con los de la red

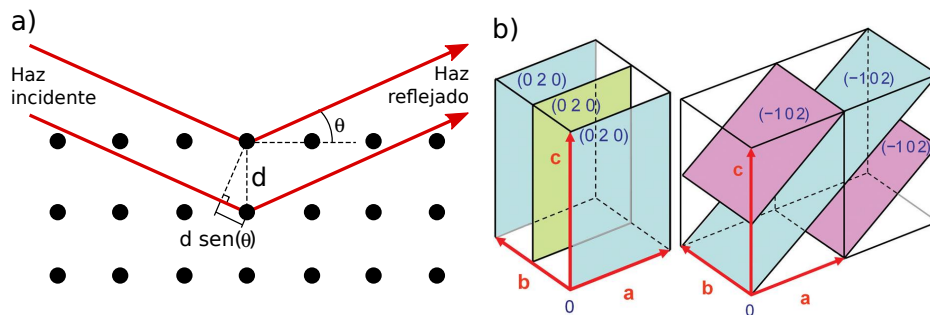
directa (Figura 3.13a). Los puntos de la red recíproca se caracterizan por tres valores enteros denominados *índices de Miller* o, de forma genérica, índices  $hkl$  (Figura 3.13b). Este artificio matemático es de especial utilidad en cristalografía ya que el patrón de difracción obtenido al final del experimento se corresponde con la red recíproca del cristal. Así pues, en términos generales, conociendo las herramientas para construir una red a partir de la otra y habiendo obtenido de forma experimental la red recíproca del cristal, se puede reconstruir la red directa asociada y, con ello, las coordenadas de cada átomo en el espacio vectorial real.



**Figura 3.13: Red directa y red recíproca.** a) Red recíproca bidimensional (azul) calculada a partir de una red directa bidimensional (roja). b) Red recíproca tridimensional, mostrando los índices de Miller de cada nodo. Figura adaptada de Rupp, 2009.

### 3.3.2.3. Interacción geométrica de la luz con el cristal

Al llegar los rayos X al cristal, sufren una difracción que genera varios haces dispersados en diferentes direcciones y con diferente intensidad. Se puede, por tanto, describir un modelo en el que, siguiendo la Ley de Bragg (Bragg, 1913), los átomos del interior del cristal se agrupan en planos que reflejan el haz incidente hacia una dirección determinada (Figura 3.14a). Estos planos reciben el nombre de *planos de Miller* o *planos cristalográficos*, caracterizados por los índices de Miller correspondientes a los recíprocos de las intersecciones de dichos planos con los ejes  $xyz$  de la red directa (Figura 3.14b).

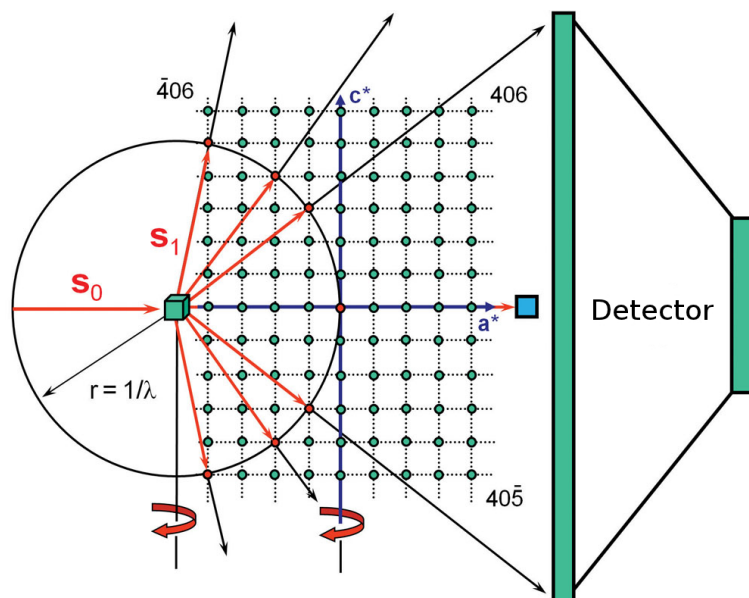


**Figura 3.14: Ley de Bragg y planos de Miller.** a) Esquema de la Ley de Bragg, describiendo la reflexión de un haz de luz por una serie de planos formados por átomos. Su expresión es  $n\lambda = 2d \sin(\theta)$ , donde  $\lambda$  es la longitud de onda del haz incidente y  $n$  un número entero. b) Representación en el espacio real de los planos de Miller, indicando sus índices  $hkl$ . Figura adaptada de Rupp, 2009.

### 3.3.2.4. Esfera de Ewald y registro del patrón de difracción

Visto lo anterior, cada reflexión que surge de la difracción de los rayos X está asociada a un plano de Miller determinado. Por tanto, las reflexiones se describen con los mismos índices  $hkl$  que los del plano al que se asocian, que a su vez se corresponden con los nodos de la red directa asociada al cristal. Así pues, se puede determinar que cada reflexión recogida en el patrón de difracción aporta la información de un punto de la red recíproca, de coordenadas  $(h, k, l)$  dentro del espacio recíproco.

Para explicar matemáticamente la relación entre la radiación incidente, las reflexiones generadas y su correspondencia con los puntos de la red recíproca, se utiliza una construcción geométrica llamada *esfera de Ewald* (Ewald, 1969). Esta herramienta permite, en función de la longitud de onda del haz incidente y de la inclinación de los planos de Miller en cada momento durante la rotación del cristal, describir la disposición espacial de los puntos del patrón de difracción (Figura 3.15).



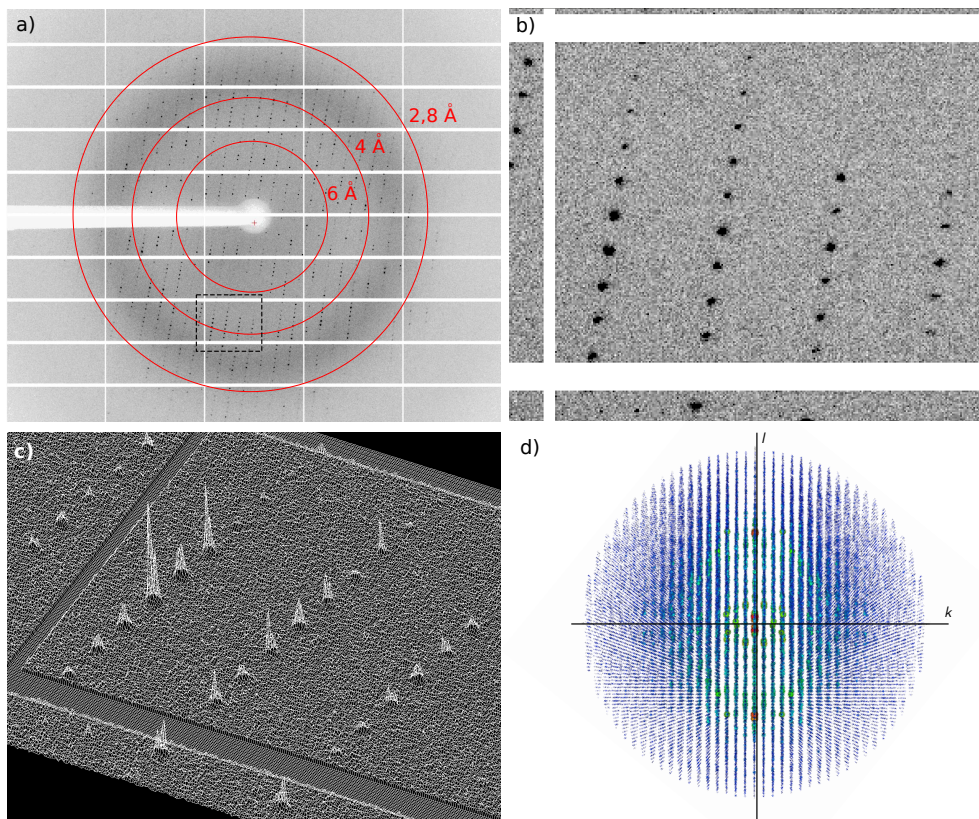
**Figura 3.15: Esfera de Ewald.** La esfera de Ewald, centrada en el cristal (cubo verde) y de radio  $1/\lambda$ , sirve para describir las condiciones en las que se generan los haces difractados. El haz incidente ( $s_0$ ) se difracta en diferentes haces ( $s_1$ ) cuando, durante la rotación del cristal, un punto de la red recíproca interseca con la superficie de la esfera. Dicho haz tomará la dirección del cristal a dicho punto y será registrado en el detector. Los parámetros  $a^*$  y  $c^*$  hacen referencia a dos de los ejes de la red recíproca, cuyo centro se sitúa sobre la superficie de la esfera. Figura adaptada de Rupp, 2009.

Los rayos X reflejados en diferentes direcciones son recogidos por el detector durante el tiempo que dure el experimento. Al rotar el cristal rota la red directa y, por tanto, también rota la red recíproca asociada, haciendo que sus puntos entren y salgan de la esfera de Ewald progresivamente. Por tanto, la técnica es capaz de registrar gran parte de la red recíproca del cristal en tres dimensiones. Además, como se ha dicho anteriormente, este modelo sirve de puente entre la red recíproca y la red directa, hecho especialmente interesante durante el análisis de los datos de difracción y la geometría del cristal.



### 3.3.2.5. Información aportada por el patrón de difracción

Ya que las reflexiones se registran progresivamente a medida que el cristal rota, el patrón de difracción final se compone de una serie de cortes bidimensionales del registro tridimensional de la red recíproca (Figura 3.16). Cada punto del patrón de difracción se corresponde con un punto de la red recíproca.



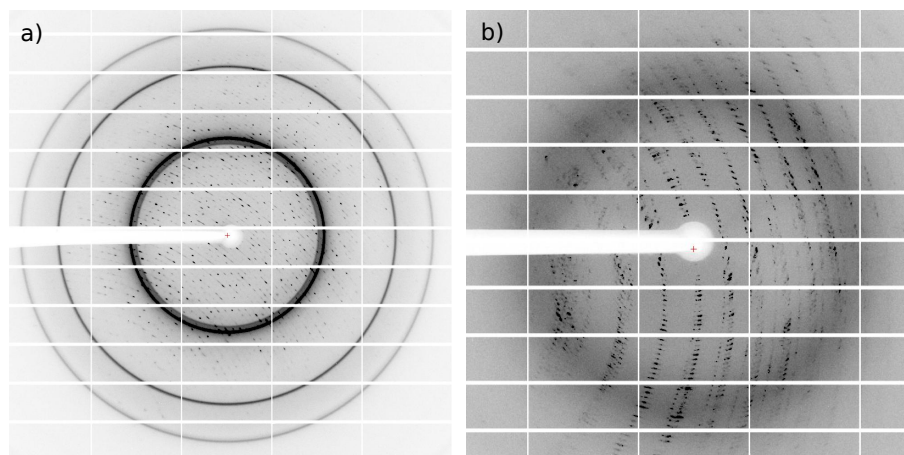
**Figura 3.16: Patrón de difracción.** **a)** Patrón de difracción de la proteína DacB (resuelta en este trabajo), medido con Pilatus 6M (Dectris), mostrando los anillos de resolución. La cuadrícula blanca se corresponde con zonas ciegas del detector. La sombra blanca se debe al bloqueador del haz, que evita que la zona central del detector se dañe debido a la alta intensidad de la fracción del haz incidente no dispersado. **b)** Detalle del patrón del panel **a**, correspondiente a la región recuadrada, mostrando reflexiones de distinta intensidad. **c)** Representación tridimensional de la intensidad de las reflexiones mostradas en el panel **b**. **d)** Representación tridimensional de la red recíproca medida en el patrón del panel **a**. Las reflexiones exteriores (alta resolución) son más débiles que las interiores (baja resolución). El eje  $h$  aparece orientado hacia el observador.

La información básica que contienen los puntos del patrón y que se utiliza para la posterior resolución estructural es la siguiente:

- **Posición:** Dada por las coordenadas del espacio recíproco  $(h, k, l)$ .
- **Tamaño y forma:** Informan de la intensidad de la reflexión que ha generado cada punto, representada por  $I_{hkl}$ .

En función de la simetría del cristal y de otros parámetros, como la mosaicidad o la resistencia a la radiación, se debe planificar una estrategia de recogida de datos que sea adecuada para cada caso particular (Dauter, 1999; Evans, 1999; Pflugrath, 1999; Dauter, 2005; Bourenkov &

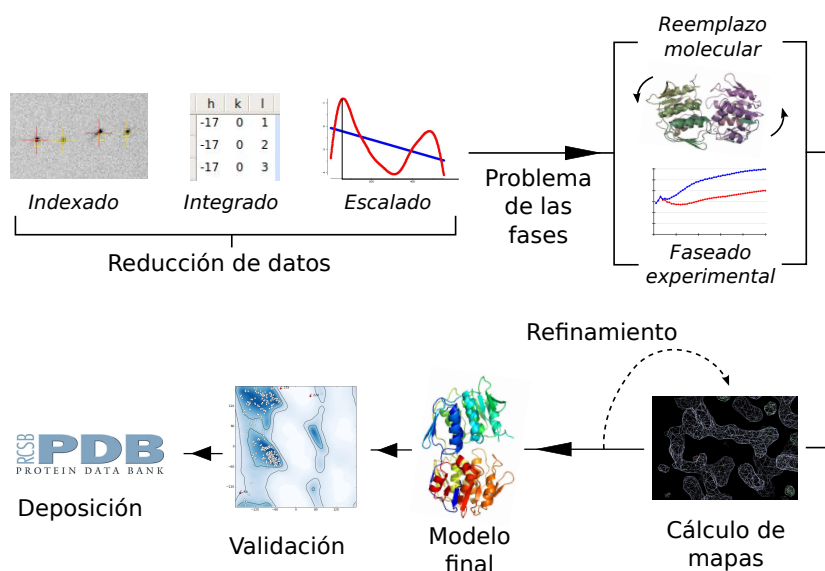
Popov, 2006). Además, la inspección visual del patrón de difracción ayuda de forma indirecta a determinar ciertas propiedades del experimento, como la resolución alcanzada, la simetría del espacio recíproco, la presencia de hielo en la muestra, presencia de maclas, daño por radiación, etcétera (Figura 3.17).



**Figura 3.17: Defectos en los patrones de difracción.** a) Patrón de un cristal de la proteína DacB (resuelta en este trabajo) sobre el que se ha formado hielo (anillos concéntricos). b) Patrón de difracción de un cristal de la proteína DacB excesivamente dañado, con puntos alargados debido a una mosaicidad extrema y fracturas del cristal.

### 3.3.3. Determinación de la estructura tridimensional de la proteína

Una vez que se ha obtenido el patrón de difracción experimental, se suceden varias etapas para analizar la información y extraer la estructura tridimensional de la proteína. En la Figura 3.18 se resumen las diferentes etapas que serán brevemente explicadas en las sucesivas secciones.



**Figura 3.18: Etapas del procesamiento de un patrón de difracción.** En primer lugar, se deben analizar los datos y procesarlos para homogeneizar la información adquirida durante el experimento de difracción. Posteriormente, se debe solucionar el problema de las fases (principalmente mediante reemplazo molecular o faseado experimental) para después calcular los primeros mapas. Tras varios ciclos de refinamiento se obtiene un modelo que, si pasa la validación, estará listo para depositar en las bases de datos públicas.

El objetivo final es obtener la densidad electrónica de la proteína,  $\rho(x, y, z)$ , una función tridimensional que se corresponde con la envolvente producida por los electrones que rodean cada átomo de la proteína. Se describe mediante

$$\rho(x, y, z) = \frac{1}{V} \sum_{-h}^h \sum_{-k}^k \sum_{-l}^l |F_{hkl}| \cdot e^{-2\pi i(hx + ky + lz - \varphi_{hkl})} \quad (3.1)$$

donde las únicas incógnitas son  $|F_{hkl}|$  y  $\varphi_{hkl}$ . En las siguientes secciones se describe el significado y el procedimiento para calcular estos parámetros y, por tanto, la densidad electrónica de la proteína que compone el cristal.

### 3.3.3.1. Reducción de datos y cálculo de $|F_{hkl}|$

En primer lugar, se utilizan algunas imágenes del patrón de difracción para determinar las propiedades de la red recíproca a partir de las posiciones de las reflexiones sobre el patrón, proceso que se conoce como *indexado*. Con esta información se pueden calcular las dimensiones de la celdilla unidad ( $a, b, c, \alpha, \beta, \gamma$ ) y estimar el posible grupo espacial característico del cristal. También se hace una estimación de la mosaicidad (ver Sección 3.2.6), que será optimizada en la siguiente etapa de la reducción de datos.

Hay que tener en cuenta que, al registrar una rotación de un cristal que contiene diferentes elementos de simetría, cada reflexión puede medirse varias veces en distintos momentos del experimento. Y además, las imágenes pueden producir cortes en los puntos, de modo que una misma reflexión aparece cortada y registrada en imágenes consecutivas. Por tanto, tras el indexado se realiza una etapa de *integrado* consistente en calcular la posición ( $hkl$ ) e intensidad ( $I_{hkl}$ ) de cada reflexión a lo largo de todas las imágenes registradas, agrupando la información de las múltiples veces que ha sido registrada a lo largo de la rotación del cristal (Leslie, 2006). En este punto surge un parámetro importante en el procesamiento de datos, denominado *multiplicidad* que se corresponde con el número de veces que se ha registrado la misma información. A mayor multiplicidad, más precisa será la estimación de la intensidad para una reflexión  $hkl$  determinada. Además, también es importante medir el mayor porcentaje de espacio recíproco posible, lo que se representa mediante un segundo parámetro denominado *completitud*. Valores por encima del 90 % para la totalidad del patrón en términos de resolución son más que aceptables, mientras que en la capa de más alta resolución, donde se agrupan las reflexiones más débiles, se permite una completitud ligeramente inferior al 90 %.

Debido a posibles defectos inhomogéneos del cristal o al daño por radiación, no se puede considerar que los datos tomados al inicio del experimento o en determinada zona del cristal vayan a presentar la misma calidad que los tomados al final del experimento o en otra zona de incidencia del haz durante la rotación. Por tanto, debido a esta falta de correlación, se hace necesaria una etapa de *escalado*, que enfrente los datos recogidos en cada momento del experimento con parámetros como el decaimiento de la difracción por daño por radiación, la mosaicidad, etcétera (Evans, 2006).

Una vez que los datos se han indexado, integrado y escalado, se dispone de información objetiva y normalizada que permite analizar la calidad de los datos tanto de forma global como



por capa de resolución (Rossmann & van Beek, 1999; Kabsch, 2010a; Evans, 2011). Además de la multiplicidad y la completitud, introducidos anteriormente, existen algunos parámetros imprescindibles a la hora de determinar el límite de resolución del patrón de difracción y su calidad general (Weiss, 2001; Rupp, 2009; Karplus & Diederichs, 2012; Evans & Murshudov, 2013). A continuación se describen los indicadores más utilizados durante este trabajo para el análisis de los datos de difracción:

- $\langle |I|/\sigma(I) \rangle$ : La intensidad de las reflexiones disminuye a medida que nos acercamos a las capas de mayor resolución, donde las reflexiones son especialmente débiles. Así pues, el patrón se divide en capas de resolución para las que se calcula el valor de  $\langle |I|/\sigma(I) \rangle$ , cuya expresión es

$$\langle |I|/\sigma(I) \rangle = \frac{1}{N} \sum_{hkl} \frac{|I_{hkl}|}{\sigma(I_{hkl})} \quad (3.2)$$

y representa cuántas veces es más intensa la señal de la reflexión que el ruido de fondo. Así pues, se suele describir este indicador como una relación señal/ruido. Un  $\langle |I|/\sigma(I) \rangle \simeq 2$  se considera un valor aceptable para poder discriminar las reflexiones del ruido, aunque con los detectores más modernos puede utilizarse un límite algo menor. El valor  $N$  hace referencia al número de reflexiones de la capa de resolución para la que se ha calculado  $\langle |I|/\sigma(I) \rangle$ .

- $R_{\text{merge}}$ : Este valor es un factor estadístico de desacuerdo entre las intensidades de las reflexiones relacionadas por simetría dentro del patrón de difracción (Arndt *et al.*, 1968). Nuevamente, se puede calcular un valor global o diferenciar entre capas de resolución, para detectar anomalías en determinadas zonas del patrón. La expresión de este indicador estadístico es

$$R_{\text{merge}} = \frac{\sum_{hkl} \sum_{i=1}^N |I_{hkl,i} - \bar{I}_{hkl}|}{\sum_{hkl} \sum_{i=1}^N I_{hkl,i}} \quad (3.3)$$

donde  $\bar{I}_{hkl}$  es la intensidad media de cada reflexión. Un valor bajo en este estadístico indica un buen acuerdo entre las diferentes medidas de cada reflexión durante todo el experimento. No obstante, los errores cometidos en cada medida de una determinada reflexión se acumulan. Así pues, si la multiplicidad es muy elevada (algo habitual con la tecnología actual), la medida del error puede aumentar considerablemente, elevando el  $R_{\text{merge}}$  por encima de los valores aceptables. Por este motivo, este indicador puede ser fiable para reflexiones muy intensas (con bajo error asociado), pero no debe ser considerado a la hora de determinar el corte de la alta resolución del conjunto de datos (Evans, 2011; Evans & Murshudov, 2013).

- $R_{\text{pim}}$ : Ya que una mayor multiplicidad es sinónimo de una mayor precisión en la medida de una reflexión determinada, es necesario incluir una modificación en el  $R_{\text{merge}}$  para que no se comporte de manera opuesta a lo adecuado (Diederichs & Karplus, 1997). Esa modificación pasa por normalizar el efecto de  $N$ , generando un nuevo indicador, denominado  $R_{\text{pim}}$

(*precision-indicating merging R-value*). Su expresión es, por tanto, similar a la del  $R_{\text{merge}}$ , siendo la única diferencia la inclusión de la componente normalizadora que tiene en consideración el número de reflexiones medidas:

$$R_{\text{pim}} = \frac{\sum_{hkl} \left( \frac{1}{N-1} \right)^{1/2} \sum_{i=1}^N |I_{hkl,i} - \bar{I}_{hkl}|}{\sum_{hkl} \sum_{i=1}^N I_{hkl,i}} \quad (3.4)$$

El hecho de que el  $R_{\text{pim}}$  disminuya con la multiplicidad significa que tiene en consideración la mejoría en la medida de una reflexión determinada debida a la suma de varias medidas independientes. Por este motivo, el  $R_{\text{pim}}$  se considera un indicador mucho más útil que el  $R_{\text{merge}}$  para la estimación de la calidad de un patrón de difracción (Weiss, 2001; Evans, 2011). Un valor de  $R_{\text{pim}}$  por debajo de 0,3 – 0,4 es bastante aceptable.

- **$CC_{1/2}$  y  $CC^*$** : Recientemente se ha propuesto un nuevo parámetro, basado en el cálculo del coeficiente de correlación de Pearson, que determina con mayor precisión la calidad de los datos de difracción (Karplus & Diederichs, 2012). Si se divide el conjunto de datos en dos mitades, eligiendo de forma aleatoria las reflexiones que pertenecen a cada nuevo subconjunto, los errores obtenidos en cada subconjunto también serán aleatorios y de magnitud similar. El  $CC_{1/2}$  es el coeficiente de correlación entre las reflexiones de ambos subconjuntos. Esta correlación será elevada para las capas de menor resolución, por contener las reflexiones más intensas (mejor medidas), mientras que en las capas de alta resolución las reflexiones se empiezan a mezclar con el ruido y la correlación será menor. Otro indicador relacionado con el  $CC_{1/2}$  es el  $CC^*$ , que muestra la correlación real entre los datos observados y las intensidades reales desconocidas y cuya expresión viene dada por

$$CC^* = \sqrt{\frac{2CC_{1/2}}{1 + CC_{1/2}}} \quad (3.5)$$

Ambos indicadores están extendiendo su uso y se están incorporando a las estadísticas básicas de los programas de reducción de datos (Diederichs & Karplus, 2013; Evans & Murshudov, 2013).

Existen otros parámetros adicionales que nos permiten, llegados a este punto, analizar ciertos defectos en el cristal, como el grado de anisotropía o la presencia de maclas. Un análisis más detallado sobre la detección y análisis de maclas se realizará en la sección 3.3.4. Además, conociendo los parámetros básicos de la celdilla unidad y el grupo espacial se puede calcular el volumen de la unidad asimétrica. Con este dato y con el peso molecular de la proteína que forma el cristal se puede determinar el número de moléculas por unidad asimétrica, el porcentaje de solvente del cristal y el número de  $\text{\AA}^3$  por Dalton, lo que se denomina coeficiente de Matthews (Matthews, 1968).

Finalmente, se procede al cálculo del parámetro  $|F_{hkl}|$ , primera incógnita de la función de densidad electrónica (Ecuación 3.1). Agrupando todos los factores de dispersión atómica ( $f$ ) de

los átomos que componen la celdilla unidad de un cristal ( $N$ ) sobre el que se irradia un haz de rayos X, se obtiene un factor de dispersión global, denominado *factor de estructura*, cuya expresión es

$$F_{hkl} = \sum_{j=1}^N f_j e^{2\pi i(hx_j + ky_j + lz_j)} \quad (3.6)$$

Cada reflexión detectada en el patrón de difracción tiene asociado un factor de estructura  $F_{hkl}$  derivado de la dispersión del haz por determinados átomos del cristal. Por tratarse de una onda, el factor de estructura  $F_{hkl}$  se divide en módulo ( $|F_{hkl}|$ ) y fase ( $\varphi_{hkl}$ ), las dos incógnitas de la ecuación 3.1.

El módulo del factor de estructura de cada reflexión es proporcional a su intensidad, por lo que el valor de  $|F_{hkl}|$  se extrae directamente del análisis del patrón de difracción tras la etapa de escalado. En cambio, durante la recogida de datos se pierde la información de la fase del factor de estructura, dando lugar al denominado *problema de las fases*, que se desarrollará a continuación.

Durante este trabajo se han utilizado los programas XDS (Kabsch, 2010b) e iMosflm (perteneciente al conjunto de programas de CCP4, *Collaborative Computational Project No. 4*; Battye *et al.*, 2011; Winn *et al.*, 2011) para el indexado e integrado de los datos. El escalado de los datos procesados con XDS se ha realizado mediante XSCALE (incluido en el programa XDS), mientras que los datos procesados con iMosflm se escalaron con Scala (Evans, 2006) y Aimless (Evans & Murshudov, 2013). Este último programa incluye, además de una versión mejorada de Scala, el programa Pointless, útil para analizar la simetría del cristal (Evans, 2006).

### 3.3.3.2. El problema de las fases y cálculo de $\varphi_{hkl}$

El problema de las fases surge por la imposibilidad de calcular  $\varphi_{hkl}$  a partir del patrón de difracción, ya que lo que se recogen son intensidades, proporcionales al módulo del factor de estructura pero sin relación alguna con la fase. Así pues, con el experimento básico de difracción de rayos X no se puede calcular directamente la función de densidad electrónica  $\rho(x, y, z)$  (Ecuación 3.1). Por tanto, es necesario utilizar técnicas accesorias que nos faciliten unas fases iniciales aproximadas, útiles para la construcción de un modelo inicial de nuestra proteína, que serán mejoradas progresivamente hasta obtener un modelo final fiel a la realidad (Rupp, 2009; Taylor, 2010; Giacovazzo *et al.*, 2011; Giacovazzo, 2013).

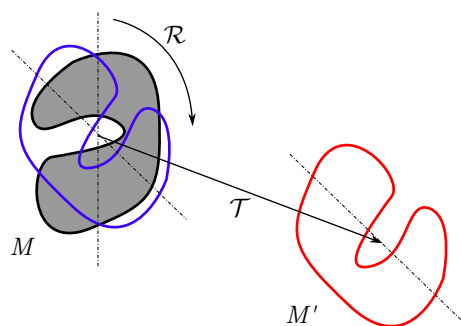
El cálculo de las fases correctas es bastante delicado, ya que si se incluyen fases erróneas en el cálculo de la densidad electrónica se producirá un sesgo que quitará peso a los factores de estructura, alterando considerablemente el modelo final hacia una solución irreal. Como se verá más adelante (Sección 5.3.4) evitar el sesgo en los mapas de densidad electrónica es de especial importancia en los casos en los que existe una macla.

Las técnicas que se utilizan para calcular las fases se agrupan en cuatro áreas, a saber, métodos directos, modificación de densidad, reemplazo molecular y faseado experimental de subestructuras (Rupp, 2009). En la práctica, los métodos directos son aplicables en datos de muy alta resolución (mejor de 1,2 Å), aunque recientemente se están haciendo avances para

poder resolver datos de menor resolución (Usón & Sheldrick, 1999; Rodríguez *et al.*, 2009). Las técnicas de modificación de densidad se suelen aplicar tras haber faseado con otro método, para mejorar las fases y los mapas resultantes (Cowtan, 2010). Las dos técnicas restantes son las más utilizadas tanto en general como en este trabajo en particular. En las siguientes secciones se realiza una breve descripción de sus fundamentos.

### 3.3.3.3. Método de reemplazo molecular

Este método es uno de los más utilizados a día de hoy, debido a la relativa sencillez que supone llevarlo a cabo y gracias a la existencia de un gran número de proteínas previamente resueltas (Rossmann & Blow, 1962; Rossmann, 2001; Scapin, 2013). La técnica consiste en utilizar como fases iniciales las obtenidas de otra proteína similar a la que se pretende resolver, presuponiendo que secuencias iguales se organizan en plegamientos iguales. Para ello, el modelo tridimensional de la proteína similar (modelo de reemplazo) se enfrenta con los datos experimentales del cristal de la proteína a resolver y, tras varios ciclos de reorganización espacial, se intenta que el modelo encaje lo mejor posible en la posición de la proteína en el cristal (Figura 3.19).



**Figura 3.19: Método de reemplazo molecular.** El reemplazo molecular se puede interpretar como una serie de traslaciones ( $\mathcal{T}$ ) y rotaciones ( $\mathcal{R}$ ) que se realizan sobre una proteína modelo ( $M$ ), seguida de una evaluación que permite discriminar si la posición obtenida ( $M'$ ) es correcta o si se debe seguir buscando otra orientación distinta. Si el modelo es realmente similar a la proteína problema, la solución del reemplazo molecular se acercará notablemente a la real.

Una vez conseguido ese paso se obtiene un primer mapa de densidad electrónica con unas fases iniciales provenientes del modelo. Como se verá más adelante (Sección 3.3.3.5), tras numerosos pasos de construcción manual y refinamiento, ajustamos las fases a nuestra proteína real y mejoramos la calidad tanto del mapa de densidad electrónica como del modelo de la proteína.

En casos complicados, con proteínas modulares o con fragmentos conservados aislados, es posible que no exista un modelo de reemplazo adecuado que cubra la totalidad de la proteína. En esos casos, el reemplazo molecular no es en absoluto trivial y requiere un trabajo previo de búsqueda y preparación de un modelo de reemplazo adecuado, cuya dificultad puede variar en función de la información de la que disponemos, el porcentaje de similitud de nuestra proteína con otras, la separación entre dominios o zonas conservadas, etcétera (Bunkóczi & Read, 2011). Generalmente, la fabricación del modelo de reemplazo requiere la fusión de diferentes dominios en una sola proteína o la búsqueda independiente de cada dominio dentro del cristal, con la esperanza de obtener unas fases preliminares que sirvan para predecir la densidad electrónica de las partes de nuestra proteína que no están contenidas en el modelo de reemplazo.

Los indicadores básicos para determinar si una solución de reemplazo molecular es correcta son ciertas variables normalizadas (*Z-scores*) que se aplican a la función de rotación (RFZ, *Rotation Function Z-score*) y a la función de traslación (TFZ, *Translation Function Z-score*) durante el cálculo de orientación de la estructura modelo. Valores de TFZ y RFZ superiores a 7 indican una buena probabilidad de haber encontrado la solución correcta. Además, se utilizan otros parámetros adicionales para evaluar los choques por empaquetamiento.

Los programas más utilizados en este trabajo para realizar el reemplazo molecular han sido Molrep (Vagin & Teplyakov, 2010) y Phaser (McCoy *et al.*, 2007), ambos disponibles en el conjunto CCP4. Además, en alguna ocasión se han utilizado otras aplicaciones que automatizan el proceso de búsqueda del modelo óptimo, el reemplazo molecular propiamente dicho y un refinamiento posterior para mejorar los mapas, como es el caso de MrBUMP (Keegan & Winn, 2008) y Balbes (Long *et al.*, 2008). Estos programas permiten encontrar soluciones complejas en un tiempo mucho menor que el que se requeriría de forma manual, ya que son capaces de ensayar múltiples modelos en múltiples posiciones en una única ejecución.

Otro programa de especial utilidad en el reemplazo molecular es Chainsaw (Stein, 2008), perteneciente al conjunto CCP4. Ya que los modelos utilizados para el reemplazo molecular suelen pertenecer a otros microorganismos, su secuencia peptídica suele diferir en un alto grado de la secuencia de la proteína que se intenta construir. Por ello, tras realizar un reemplazo molecular con un modelo distinto a la proteína objetivo, es bastante habitual que la cadena principal aparezca bien ubicada en la densidad electrónica mientras que las cadenas laterales no se corresponden con las correctas. El programa Chainsaw permite mutar la secuencia original por la final, generando un modelo de reemplazo molecular cuya secuencia se corresponde con la secuencia de la proteína objetivo. Así pues, tras el reemplazo molecular, no sólo la cadena principal aparece bien ubicada, sino que las cadenas laterales son las correctas y sólo necesitarán ser reorientadas hacia la densidad electrónica observada. Este programa además sirve para construir modelos de poli-Ala, útiles en casos de baja resolución o con sesgo.

#### 3.3.3.4. Faseado experimental de subestructuras

Cuando no se dispone de una proteína previamente resuelta similar a la que queremos resolver, se requieren otras técnicas que son capaces de obtener las fases reales de nuestra proteína a partir de experimentos de difracción adicionales, sin requerir el apoyo de fases similares ficticias. La base de todas las técnicas de faseado mediante el uso de subestructuras es común y consiste en introducir átomos en el cristal que modifiquen el patrón de difracción de tal forma que se pueda calcular primero una subestructura formada por dichos átomos y, posteriormente, tomando esta como referencia, la estructura de la proteína (Giacovazzo *et al.*, 2011; Giacovazzo, 2013).

Las técnicas que hacen uso de estos fundamentos se pueden agrupar en dos grupos:

- **Técnicas de reemplazo isomorfo simple/múltiple (SIR/MIR):** Se introducen átomos de elevado número atómico en el cristal que alterarán la dispersión de los rayos X, con la intención de generar un patrón de difracción que, comparado con el cristal sin estos átomos, permita estimar unas fases iniciales. Es requisito indispensable que se mantenga el isomorfismo entre el cristal nativo (original) y el cristal derivado (que contiene los átomos

pesados), para que la comparación de ambos patrones sea eficaz. Si se utiliza un sólo cristal, se habla de un reemplazo isomorfo simple (SIR, *Single Isomorphous Replacement*), mientras que si se utilizan diferentes cristales derivados se habla de un reemplazo isomorfo múltiple (MIR, *Multiple Isomorphous Replacement*). Estas técnicas fueron las primeras utilizadas para resolver estructuras de proteínas (Robertson, 1935; Green *et al.*, 1954).

- **Técnicas de dispersión anómala simple/múltiple (SAD/MAD):** Los átomos que se introducen en el cristal resuenan al ser irradiados a determinada longitud de onda, lo que genera una dispersión anómala detectable en el patrón de difracción. Las diferencias de intensidad detectadas producidas por este efecto se utilizan para estimar las fases iniciales. Si se utiliza una única longitud de onda durante la medida de los datos se habla de dispersión anómala simple (SAD, *Single-wavelength Anomalous Dispersion*) mientras que si se miden diferentes patrones a diferentes longitudes de onda se utiliza la dispersión anómala múltiple (MAD, *Multi-wavelength Anomalous Dispersion*). Estas técnicas empezaron a aplicarse varios años después de la aplicación del reemplazo isomorfo (Karle & Karle, 1981; Hendrickson, 1991).

Como se ha indicado previamente, la preparación de un experimento de este tipo implica la introducción de átomos adicionales en el cristal (Garman & Murray, 2003). Además, estos átomos deben estar ordenados para que participen en la dispersión de los rayos X. Las técnicas más utilizadas para introducir átomos pesados en un cristal de proteína son las siguientes:

- **Sustitución:** Se sustituyen algunos átomos de la proteína por átomos pesados. El caso más habitual es la sustitución de metioninas o cisteínas por selenometioninas (Se-Met) o selenocisteínas (Se-Cys), donde el átomo de azufre ha sido sustituido por un átomo de selenio, fuerte dispersor anómalo.
- **Inmersión:** Al igual que en el caso de la cristalización en presencia de ligandos (Sección 3.2.3), se puede sumergir un cristal en una solución que contenga los átomos adicionales disueltos, esperando que penetren en el cristal mediante los canales de solvente característicos de los cristales de proteína. Los átomos adicionales pueden anclarse a la estructura de manera fija y ordenada mediante enlaces con las cadenas laterales de algunos residuos. Las concentraciones y los tiempos de incubación varían dependiendo de cada caso, por lo que conviene realizar diferentes ensayos con diferentes átomos pesados.

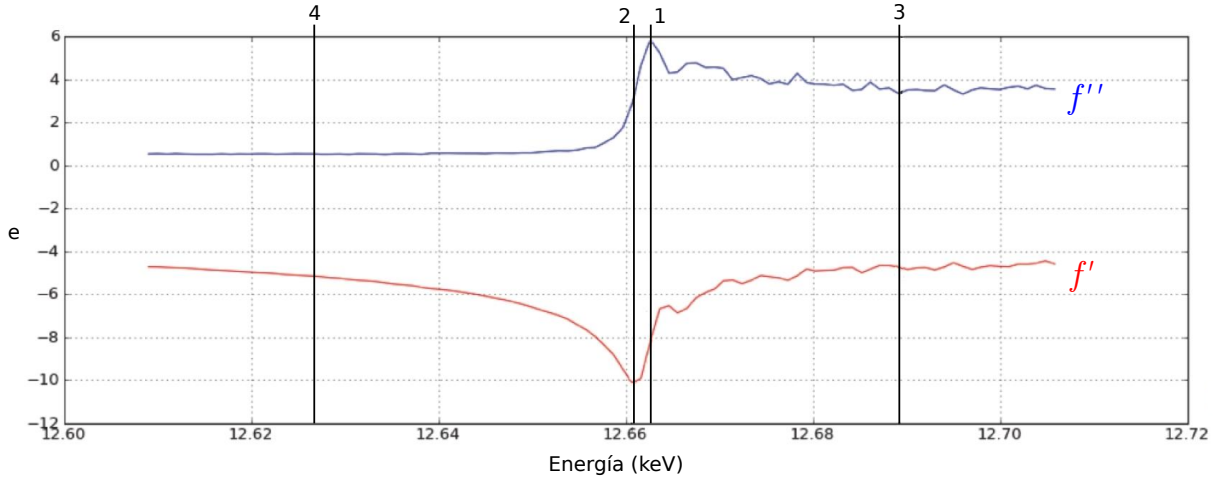
En este trabajo se han utilizado exclusivamente técnicas de dispersión anómala (SAD y MAD) como métodos para fasear mediante subestructuras. En un experimento de tipo SAD o MAD, el factor de dispersión atómico ( $f$ ) sufre una división en dos componentes, normal ( $f_0$ ) y anómala ( $f'$  y  $f''$ ), debido a la resonancia del átomo a determinada longitud de onda. La componente anómala presenta a su vez una parte real y una parte imaginaria, de modo que el factor de dispersión atómico final se puede representar como

$$f(\theta, \lambda) = f_0(\theta) + f'(\lambda) + if''(\lambda) \quad (3.7)$$

siendo  $f_0$  el factor de dispersión atómico en ausencia de dispersión anómala. La medida de datos de difracción a diferentes longitudes de onda (una de ellas la de resonancia) permite obtener unas

diferencias en los factores de estructura que llevarán, finalmente, a estimar unas fases iniciales para el cristal (Dauter *et al.*, 2002; Dodson, 2003; Giacovazzo, 2013).

Para maximizar estas diferencias anómalas, se deben elegir unas longitudes de onda adecuadas durante la medida del patrón de difracción (Figura 3.20). En primer lugar, se debe medir un conjunto de datos a la longitud de onda coincidente con la máxima capacidad dispersiva del átomo dispersor anómalo en cuestión (valor máximo de  $f''$ ; **1** en la Figura 3.20). Se puede medir otro conjunto de datos a la longitud de onda coincidente con el punto de inflexión de la curva  $f''$ , equivalente al punto en el que  $f'$  es mínimo (**2** en la Figura 3.20), lo que permite maximizar las diferencias dispersivas entre longitudes de onda. Adicionalmente se suelen medir conjuntos de datos a longitudes de onda alejadas de esta zona tanto hacia un extremo como hacia el otro (**3** y **4** en la Figura 3.20), con la intención de aumentar el contraste de la dispersión respecto a los datos tomados en la inflexión de  $f''$ .



**Figura 3.20: Espectro de fluorescencia para la elección de la longitud de onda.** En un experimento de tipo SAD/MAD, la elección de la longitud de onda se realiza en base a los máximos y mínimos de las curvas  $f'$  y  $f''$  características del átomo dispersor anómalo, calculadas en torno a su límite de absorción. En un experimento de tipo MAD se suelen elegir entre 2 y 4 longitudes de onda, correspondientes al pico (**1**) e inflexión (**2**) de la curva  $f''$  y a dos longitudes de onda remotas (**3**) y (**4**). En los experimentos de tipo SAD se mide un único conjunto de datos, generalmente a la longitud de onda del pico (**1**).

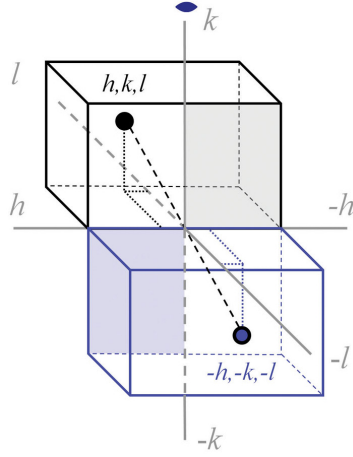
En un caso de difracción no anómala, pares de reflexiones simétricas presentan la misma intensidad y fase opuesta, hecho que se conoce como Ley de Friedel (Friedel, 1913) (Figura 3.21):

$$|F_{hkl}| = |F_{\bar{h}\bar{k}\bar{l}}| \quad \varphi_{hkl} = -\varphi_{\bar{h}\bar{k}\bar{l}}$$

En un experimento de dispersión anómala, las reflexiones relacionadas por simetría presentan módulos y fases distintas, por lo que surge una violación de la Ley de Friedel:

$$|F_{hkl}| \neq |F_{\bar{h}\bar{k}\bar{l}}| \quad \varphi_{hkl} \neq -\varphi_{\bar{h}\bar{k}\bar{l}}$$

Este hecho es el que permite calcular las diferencias anómalas entre reflexiones y utilizarlas para el faseado del patrón de difracción (Giacovazzo *et al.*, 2011; Giacovazzo, 2013).



**Figura 3.21: Ley de Friedel.** Las reflexiones relacionadas por una operación de inversión presentan la misma intensidad y fase opuesta. Figura extraída de Rupp, 2009.

La calidad de la señal anómala que obtengamos dependerá, entre otros factores, del número de dispersores anómalos unidos al cristal y de su capacidad dispersiva en el límite de absorción, por lo que dependiendo de cada caso podemos esperar mayores o menores diferencias anómalas que favorezcan en mayor o menor medida el cálculo de las fases iniciales (Boggon & Shapiro, 2000). Así pues, existen ciertos átomos que presentan una alta capacidad de dispersión anómala, como es el caso del selenio, oro, gadolinio o mercurio, ampliamente utilizados ya que un pequeño número de ellos pueden ser suficientes para fasear (como ejemplo, se estima que es necesario un átomo de selenio por cada 100 aminoácidos para realizar Se-SAD; Strub *et al.*, 2003). Otros átomos, en cambio, no son tan buenos dispersores, como el azufre (Yang *et al.*, 2003; Cianci *et al.*, 2008), que requiere un ajuste fino de la longitud de onda, la dosis y la multiplicidad para poder extraer las diferencias anómalas. Por último, también influye la longitud de onda a la que resuena cada átomo, ya que es necesario sintonizar la fuente de rayos X para que emita con la energía necesaria y, en ocasiones, esta se sitúa cerca de los límites de las líneas experimentales de radiación sincrotrón (habitualmente entre 5 y 15 keV). Un ejemplo clásico es el yodo, átomo con elevada capacidad dispersiva pero a energías muy bajas, cercanas a 5 keV (Evans *et al.*, 2003; Abendroth *et al.*, 2011).

Una vez que se han obtenido los patrones de difracción a las longitudes de onda adecuadas, se realiza un indexado, integrado y escalado sobre cada uno de ellos. Posteriormente, se tiene en cuenta la información anómala para calcular la calidad de la señal anómala, estimar el número de dispersores por celdilla unidad y determinar una subestructura. Si las diferencias anómalas son suficientemente elevadas, se podrán calcular las fases iniciales y los primeros mapas de densidad electrónica.

Un indicador habitualmente utilizado tras el escalado para estimar la calidad de los datos anómalos es el  $R_{\text{anom}}$ , cuya expresión viene dada por

$$R_{\text{anom}} = \frac{\sum_{hkl} |I_{h\bar{k}\bar{l}} - I_{hkl}|}{\sum_{hkl} \bar{I}_{hkl}} \quad (3.8)$$



siendo  $\bar{I}_{hkl}$  la intensidad media de cada reflexión  $hkl$ . Cuando el  $R_{\text{pim}}$  (Ecuación 3.4) es 1,5 veces menor que el  $R_{\text{anom}}$ , se considera que la estructura es resoluble mediante faseado anómalo (Weiss, 2001; Weiss *et al.*, 2001).

Otro parámetro fundamental para determinar la presencia de señal anómala en un conjunto de datos es el  $\langle d''/\text{sig} \rangle$ , que evalúa la relación entre señal anómala y ruido (Figura 3.22a). Su expresión es

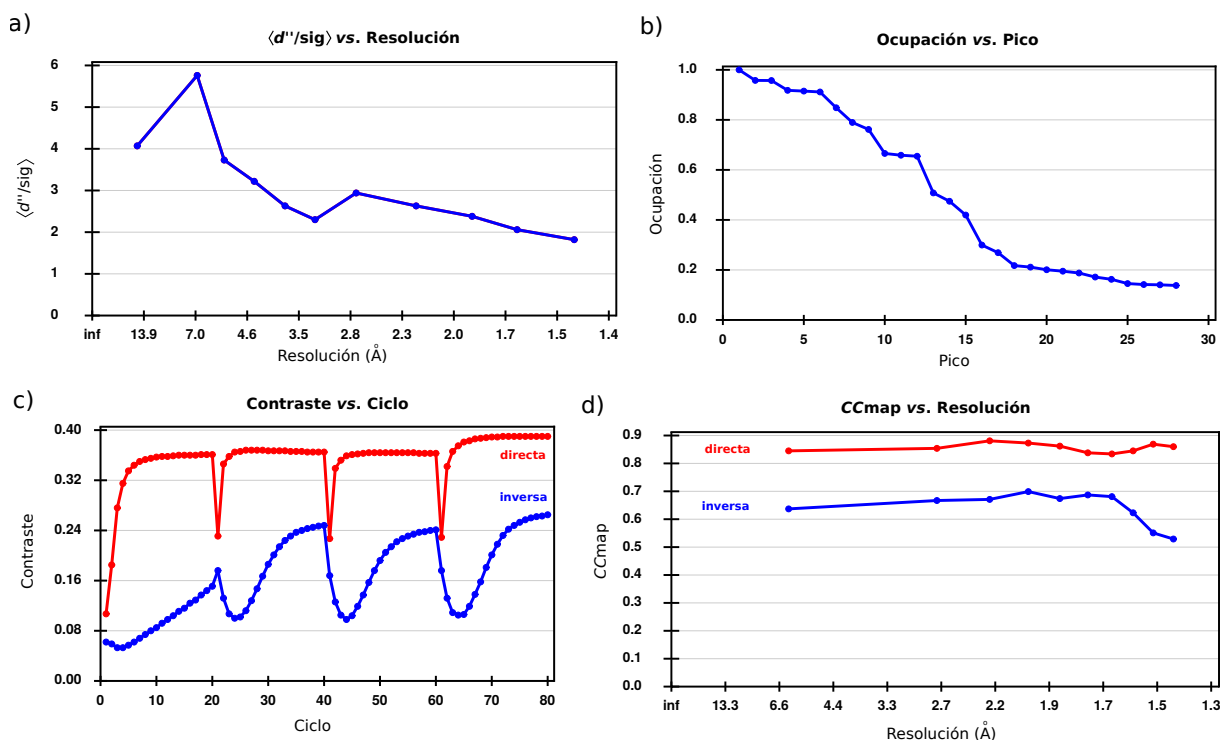
$$\langle d''/\text{sig} \rangle = \left\langle \frac{|F_{hkl} - F_{\bar{h}\bar{k}\bar{l}}|}{\sigma(F_{hkl} - F_{\bar{h}\bar{k}\bar{l}})} \right\rangle \quad (3.9)$$

El parámetro  $\langle d''/\text{sig} \rangle$  se calcula para cada capa de resolución y su valor aumenta a medida que esta disminuye. Un valor de  $\langle d''/\text{sig} \rangle$  superior a 1,3 indica la presencia de señal anómala, si bien cuanto más alto sea este valor ( $\simeq 2, 3, \dots$ ) mayor serán las diferencias anómalas en la capa de resolución considerada. Además de este valor, actualmente se suele considerar un coeficiente de correlación ( $CC_{\text{anom}}$ ) que indica a partir de qué capa de resolución la señal anómala comienza a ser significativa ( $CC_{\text{anom}} > 30\%$ ).

Durante la búsqueda de la subestructura de átomos pesados (Figura 3.22b), el procedimiento habitual consiste en situar dichos átomos aleatoriamente en el espacio, calculando unos coeficientes de correlación que relacionan la subestructura con los datos de difracción. Para alcanzar una subestructura precisa se realizan varias iteraciones, calculando los coeficientes de correlación en cada una de ellas. De esta forma, la solución correcta será la que mayor correlación presente con los datos experimentales. Se utilizan dos coeficientes de correlación distintos, uno para la totalidad de las reflexiones ( $CC_{\text{all}}$ ) y otro para las reflexiones débiles ( $CC_{\text{weak}}$ ). Valores de  $CC_{\text{all}} \geq 30$  y  $CC_{\text{weak}} \geq 15$  empiezan a ser significativos para un experimento de tipo SAD, mientras que para un MAD deben ser algo superiores. Por último, se puede calcular una figura de mérito (CFOM) consistente en la suma de los parámetros  $CC_{\text{all}}$  y  $CC_{\text{weak}}$ , siendo su valor significativo a partir de 50.

Por último, se requiere el uso de un coeficiente de correlación de mapa ( $CC_{\text{map}}$ ) o de un parámetro de contraste que permita distinguir la subestructura obtenida de la subestructura enantiomérica (lo que se conoce como subestructuras directa e inversa). Los valores de estos parámetros vienen dados en porcentaje, siendo la subestructura correcta la que presente un contraste o  $CC_{\text{map}}$  significativamente superior al de la otra subestructura (Figuras 3.22c y 3.22d).

Durante este trabajo se han usado los programas SHELXC/D/E (Sheldrick, 2010) para el análisis de la señal anómala, la estimación de los dispersores anómalos, el faseado y la modificación de densidad, tanto de forma individual (en CCP4) como agrupados en la interfaz HKL2MAP (Pape & Schneider, 2004). En ocasiones se han utilizado otros programas orientados a etapas intermedias, como HySS para estimar el número de dispersores anómalos (Grosse-Kunstleve & Adams, 2003; del conjunto de programas PHENIX; Adams *et al.*, 2010), BP<sub>3</sub> para calcular las fases y refinar la subestructura (Pannu *et al.*, 2003; Pannu & Read, 2004) y Solomon para determinar la mano de la subestructura (Abrahams & Leslie, 1996). Por último, algunos procesamientos se realizaron con CRANK (Ness *et al.*, 2004; Pannu *et al.*, 2011), una interfaz que integra de forma automática algunos de los programas anteriormente descritos.



**Figura 3.22: Indicadores del proceso de faseado experimental.** a) Gráfica de la relación entre señal anómala y ruido ( $\langle d''/\text{sig} \rangle$ ) frente a la resolución. Se observa buena señal hasta 1,5 Å. b) Gráfica de ocupación de cada átomo pesado en relación al pico de densidad en el que se ha ubicado. Siete átomos se encuentran con ocupaciones superiores al 80 %, mientras que los siguientes picos de densidad decaen progresivamente. c) Gráfica del contraste entre la subestructura directa (rojo) y la inversa (azul), en relación con la iteración para el que se ha calculado. La subestructura directa muestra mejor contraste, por lo que previsiblemente será la solución correcta. d) Gráfica del coeficiente de correlación del mapa respecto a la resolución. Si bien en ambas subestructuras la correlación es buena a lo largo de toda la resolución, la subestructura directa (rojo) muestra nuevamente una mayor correlación que la inversa (azul).

### 3.3.3.5. Cálculo de la densidad electrónica, construcción del modelo y refinamiento

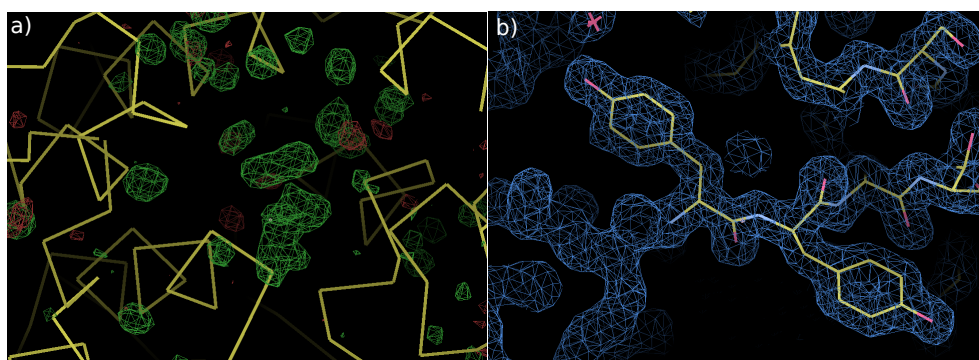
En este punto conocemos los factores de estructura experimentales y unas fases iniciales que pueden ser más o menos cercanas a las reales, dependiendo de la técnica elegida y de la calidad de los datos de difracción. Así pues, podemos calcular una primera densidad electrónica (Ecuación 3.1) que sirve de base para trazar el modelo de la proteína a resolver.

Los mapas de densidad electrónica aportan la información necesaria para construir el modelo de la proteína. Si la calidad de los mapas iniciales es muy buena, se pueden utilizar programas que intentan construir automáticamente la proteína, empezando por la cadena principal y completando las cadenas laterales progresivamente en base a la secuencia previamente conocida. Los programas más utilizados para este fin son ARP/wARP (Langer *et al.*, 2008) y Buccaneer (Cowtan, 2006). Para mapas de menor resolución o peor calidad, o para zonas de la proteína que presentan relativa flexibilidad, suele ser necesaria una construcción manual. En este aspecto, se ha utilizado exclusivamente el programa Coot (Emsley *et al.*, 2010).

Como se ha mencionado anteriormente (Sección 3.3.3.2), existen técnicas de modificación de la densidad que permiten mejorar las fases iniciales. Las técnicas más habituales se basan

en discriminar el solvente de la densidad electrónica y contrastar la frontera entre ambas zonas o en calcular una densidad media entre varias moléculas presentes en la unidad asimétrica. Los programas que se han utilizado para mejorar las fases tras un reemplazo molecular o un SAD/MAD han sido Solomon, Pirate (Cowtan, 2000) y Parrot (Cowtan, 2010).

Existen diferentes tipos de mapas que aportan la información adecuada para trazar progresivamente la estructura (Rupp, 2009). Por un lado, se suele utilizar un mapa de diferencias, conocido como  $F_{\text{obs}} - F_{\text{cal}}$ , que muestra tanto la densidad positiva (zonas que predicen átomos que aún no han sido modelados) como la negativa (zonas que han sido modeladas con átomos erróneos o que deben estar vacías). Por otro lado, también se utiliza un mapa de tipo  $2F_{\text{obs}} - F_{\text{cal}}$ , que amplifica la señal positiva y facilita la construcción en etapas iniciales del proceso (Figura 3.23).

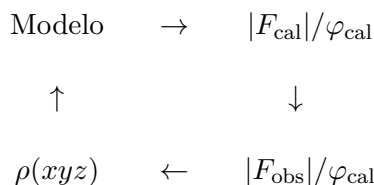


**Figura 3.23:** Mapas de densidad electrónica de uso frecuente. **a)** Mapa de la proteína MurA1 (resuelta en este trabajo), de tipo  $F_{\text{obs}} - F_{\text{cal}}$ . La densidad positiva se indica en verde y la negativa en rojo. En este caso, la densidad positiva indica moléculas de solvente aún sin modelar, mientras que la densidad negativa es despreciable. **b)** Mapa del dominio de unión a colina de la proteína CbpL (resuelto en este trabajo) de tipo  $2F_{\text{obs}} - F_{\text{cal}}$ , mostrando zonas de la proteína bien modeladas y prediciendo su continuación.

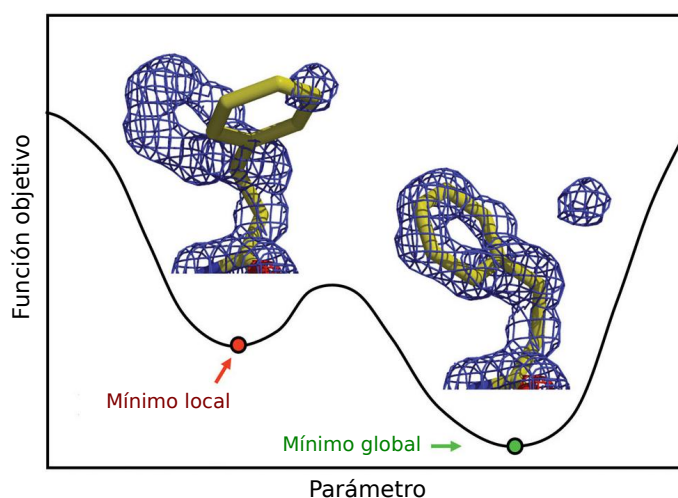
A medida que se construye el modelo, se debe ir refinando para corregir errores o desviaciones en determinados parámetros que no se pueden determinar de forma manual con precisión. El *refinamiento* es una etapa imprescindible del procesamiento de datos que consiste en enfrenar las fases del modelo nuevo con los factores de estructura experimentales, para obtener un nuevo mapa de densidad electrónica de mejor calidad que el anterior (Figura 3.24) (Rupp, 2009; Giacovazzo *et al.*, 2011).

Ya que las nuevas fases, que dan lugar a un nuevo mapa de densidad, se generan en base al modelo que se está construyendo, puede producirse un sesgo de los mapas hacia dicho modelo. Por ello, es importante que la construcción se realice de forma progresiva y sin añadir elementos que no sean predichos por los mapas. Este hecho es de especial importancia con datos maclados, como se verá en la Sección 5.3.4.

Durante el refinamiento se modifican diferentes parámetros del modelo para corregir errores menores locales, llevándolo a una situación energética más estable (Figura 3.25). A continuación se describen los parámetros más habituales que se tienen en consideración durante la etapa de refinamiento.



**Figura 3.24: Esquema del proceso de refinamiento.** Tras obtener una primera densidad electrónica  $\rho(xyz)$  se construye un modelo inicial. Con la información de este modelo se calcula su nueva densidad electrónica, lo que dará lugar a unos módulos de factores de estructura calculados ( $|F_{\text{cal}}|$ ) y unas fases calculadas ( $\varphi_{\text{cal}}$ ). Estas nuevas fases se aproximarán más a las reales que a las iniciales, mientras que los módulos de factores de estructura calculados deben ser descartados para que, en su lugar, se sigan utilizando los módulos experimentales ( $|F_{\text{obs}}|$ ). Así pues, se calcula una nueva densidad  $\rho(xyz)$ , utilizando  $|F_{\text{obs}}|$  y  $\varphi_{\text{cal}}$ , con unas fases mejoradas y la misma información experimental, lo que permitirá predecir nuevas zonas del modelo y continuar con su construcción.



**Figura 3.25: Búsqueda de mínimos globales durante el refinamiento.** La etapa de refinamiento sirve para corregir errores del modelo y llevarlo a un estado de mínima energía, pasando de mínimos locales a mínimos globales para cada parámetro a refinar. Figura adaptada de Rupp, 2009.

- **Coordenadas de cada átomo ( $x, y, z$ ):** Posiciones individuales de cada átomo en el espacio.
- **Factores de agitación térmica:** También denominados factores B, integran las desviaciones en las posiciones atómicas debidas a vibración térmica. Generalmente es un único valor que hace referencia al radio de una esfera de agitación centrada en la coordenada del átomo (refinamiento isotrópico) aunque si la resolución es muy elevada (en torno a 1,5 Å o mejor) se puede realizar un refinamiento anisotrópico que considere 6 factores B, correspondientes a los parámetros de un elipsoide de agitación. Este modelo es más realista debido a las restricciones de enlace entre diferentes átomos (Merritt, 1999; Merritt, 2012).
- **Ocupación:** Algunos átomos pueden tener una ocupación menor del 100 %, algo habitual en ligandos o moléculas del solvente, debido a una baja afinidad por la proteína. En esos casos, el refinamiento ajusta el porcentaje de ocupación a valores inferiores al 100 % (Carugo, 1999; Thouvenin & Hewat, 2000; Welsh *et al.*, 2013).

- **Simetría no cristalográfica (NCS):** En los casos en los que existe más de una molécula por unidad asimétrica, pueden aparecer relaciones de simetría entre ellas que no pertenecen a la simetría propia del grupo espacial del cristal. En estos casos, se suele hacer un refinamiento teniendo en cuenta las restricciones surgidas de estas nuevas operaciones de simetría (Vellieux & Read, 1997). Además, la NCS también puede ser útil en la etapa de reemplazo molecular o en la construcción automática del modelo, ya que se pueden encontrar o construir varias moléculas en la unidad asimétrica tomando como base una molécula previa y la NCS existente entre ellas (Tong, 2001; Isupov & Lebedev, 2008; Wiegels & Lamzin, 2012).
- **Traslación-Libración-Torsión (TLS):** Además de la agitación individual de los átomos (factores B anteriormente descritos), pueden existir movimientos de regiones particulares de la proteína a modo de cuerpo rígido. Un refinamiento de TLS ayuda a mejorar el modelo en esas zonas (Winn *et al.*, 2001; Painter & Merritt, 2006).

Además, en las etapas iniciales de la construcción y refinamiento, existen dos técnicas que pueden ayudar considerablemente a obtener un modelo fiable a partir de un modelo con regiones alejadas de los parámetros reales:

- **Enfriamiento simulado (*Simulated annealing*):** Es un método computacional que consiste en calentar la estructura a varios miles de grados e ir enfriándola paulatinamente hasta alcanzar la temperatura inicial de nuevo. Con esto se consigue ejercer una agitación que saca a la proteína de un posible mínimo de energía local y la devuelve a otro mínimo por debajo del anterior, por lo que el resultado es una estructura termodinámicamente más estable. Este método es muy útil para normalizar una solución que viene de un reemplazo molecular con un modelo muy similar al final. Si bien la técnica se conoce desde hace varios años (Kirkpatrick *et al.*, 1983; Granville *et al.*, 1994) el programa que mejor ha implementado este algoritmo recientemente es phenix.refine, perteneciente al conjunto de programas PHENIX (Adams *et al.*, 2010).
- **Cuerpo de gelatina (*Jelly-body*):** Esta opción del programa Refmac5 (Murshudov *et al.*, 2011), genera en la estructura cierta flexibilidad que permite mover grandes regiones en el espacio, lo que ayuda a reubicar dominios con relativa facilidad. Es bastante útil en el caso de proteínas modulares con regiones cuya orientación puede verse alterada bajo diferentes condiciones de cristalización.
- **Cuerpo rígido:** Si los datos de los que se dispone son de baja resolución, conviene realizar un refinamiento de cuerpo rígido en el que los átomos no se mueven individualmente, sino que toda la estructura se mueve en conjunto. De este modo, se puede ajustar en densidad la totalidad de la estructura evitando que ciertas regiones se dispersen hacia zonas de densidad débil (Delarue, 2008; Afonine *et al.*, 2009).

Además de todo lo descrito hasta el momento, los programas de refinamiento integran otras restricciones típicas en proteínas, como distancias de enlaces, ángulos, torsiones, etcétera (Evans, 2007). De esta forma, el número de parámetros a refinar disminuye considerablemente.

La forma de monitorizar el refinamiento es mediante un estadístico denominado factor  $R$  o  $R_{\text{work}}$  que mide el desacuerdo entre el modelo construido y los datos de difracción, es decir, si el modelo construido predice los datos observados (Morris *et al.*, 1992; Rupp, 2009; Giacovazzo *et al.*, 2011). Su expresión es

$$R_{\text{work}} = \frac{\sum_{hkl} ||F_{\text{obs}}| - |F_{\text{cal}}||}{\sum_{hkl} |F_{\text{obs}}|} \quad (3.10)$$

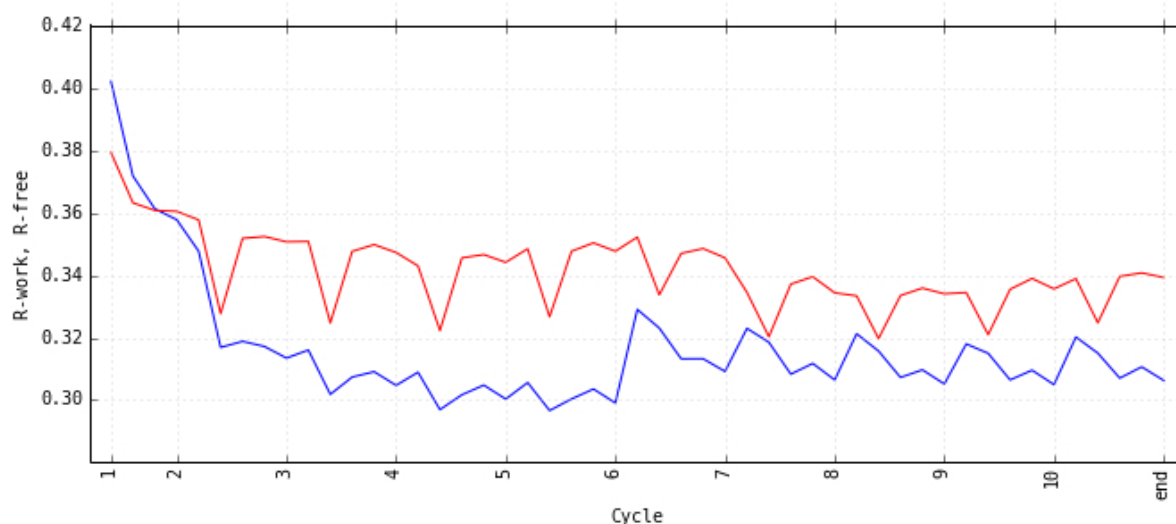
Su valor depende, en parte, de la resolución de los datos, ya que a mayor resolución, el modelo debe predecir con mayor precisión los datos de difracción, por lo que el desacuerdo tolerable es menor. En general, valores inferiores al 25 % para resoluciones intermedias son aceptables, mientras que para estructuras de alta resolución suelen alcanzar un 10-15 %.

Además del  $R_{\text{work}}$  se utiliza un segundo factor de desacuerdo, denominado  $R_{\text{free}}$  que se comporta de la misma forma que el  $R_{\text{work}}$  pero se aplica sobre un subconjunto aleatorio de reflexiones que desde el inicio del refinamiento se han mantenido separadas del resto (Brünger, 1992; Tickle *et al.*, 1998; Tickle *et al.*, 2000). De este modo, el  $R_{\text{free}}$  permite conocer si el modelo es capaz de predecir un conjunto de reflexiones que no han sido utilizadas para calcular el mapa con el que se construye el modelo. El conjunto de reflexiones que se apartan del total oscila entre un 5 y un 15 %, dependiendo de los casos.

Actualmente se está introduciendo el uso del  $CC_{\text{work}}/CC_{\text{free}}$  para analizar la correlación entre las intensidades observadas y las calculadas a partir del modelo, de manera análoga al uso del  $CC_{1/2}$  en la estimación de la calidad de los datos de difracción en contraposición a los factores  $R$  clásicos (Sección 3.3.3.1; Karplus & Diederichs, 2012; Diederichs & Karplus, 2013). No obstante, los factores  $R$  aplicados al proceso de refinamiento son indicadores más fiables que los aplicados al proceso de estimación de la calidad de los datos de difracción, por lo que su uso en esta etapa no está desaconsejado.

En la práctica, se realizan diferentes ciclos de construcción y refinamiento, mejorando progresivamente el modelo. Dependiendo de la etapa en la que se encuentre el modelo, se pueden variar los parámetros a tener en cuenta durante cada fase de refinamiento. Además, a lo largo de todo el proceso se hace un seguimiento de los factores  $R$ . Estos factores deberían disminuir a medida que se construye el modelo, ya que la predicción de los datos originales a partir del modelo mejora progresivamente. Finalmente, los factores  $R$  convergen y se estabilizan, indicando que el modelo obtenido es el mejor posible para los datos de partida (Figura 3.26).

Si se continua construyendo y refinando, con la intención de explicar más zonas de la densidad, puede producirse un fenómeno de sobre-refinamiento, caracterizado por una elevación del  $R_{\text{free}}$  tras haberse estabilizado, separándose del  $R_{\text{work}}$  en hasta más de un 10 %. Este hecho indica que el modelo está sufriendo un sesgo considerable, ya que intenta explicar más información de la que es realmente capaz. Llegados a ese punto, debe volverse sobre un modelo anterior y dar por finalizado el proceso de construcción y refinamiento de la estructura.



**Figura 3.26: Comportamiento de los factores  $R$  durante el refinamiento.** Gráfica experimental de un macrociclo de refinamiento, mostrando la disminución inicial y posterior estabilización y convergencia de los factores  $R_{\text{work}}$  (azul) y  $R_{\text{free}}$  (rojo).

Durante este trabajo se han utilizado los programas Refmac5 (Murshudov *et al.*, 2011) y phenix.refine, del conjunto de aplicaciones PHENIX (Adams *et al.*, 2010), para realizar el refinamiento de las estructuras.

### 3.3.3.6. Validación

Una vez resuelta la estructura y antes de ser depositada en la base de datos PDB, debe pasar un proceso de validación donde se analiza la geometría de todas las moléculas presentes en el cristal (Dodson, 1998; Keller *et al.*, 1998; Kleywegt, 2000). Algunos de los análisis más importantes son los relativos a los ángulos y distancias de enlace, condiciones de Ramachandran, geometría de los rotámeros de las cadenas laterales, choques entre átomos, elementos en posiciones especiales, etcétera. Además, a medida que las técnicas mejoran, se obtienen mejores procesamientos de datos que fueron previamente depositados, por lo que existe una vía de trabajo en la que se reprocesa la información antigua con la intención de corregir errores y optimizar los resultados (Joosten *et al.*, 2009).

En Coot existen herramientas que permiten analizar estos parámetros durante el proceso de construcción, útiles para evitar construir un modelo con fallos geométricos. Otros programas, como MolProbity (Chen *et al.*, 2010), realizan este tipo de validación sobre la estructura final de manera muy eficiente.

### 3.3.4. Detección, análisis y resolución de datos maclados

Como se introdujo en la Sección 3.2.6.1, las maclas son defectos de los cristales que afectan a la orientación de las moléculas y, por tanto, a la orientación del patrón de difracción resultante (Figura 3.27a). Así pues, un cristal maclado equivale, en teoría, a la suma de dos o más dominios con orientaciones diferentes (Catti & Ferraris, 1976; Chandra *et al.*, 1999; Parsons, 2003; Rupp, 2009; Giacovazzo *et al.*, 2011).

El patrón de difracción de un cristal maclado presentará la suma de las reflexiones de todos los dominios del cristal, por lo que la intensidad de cada reflexión será la suma de las diferentes contribuciones de cada dominio a dicha reflexión. Por tanto,

$$I_{hkl}^{\text{twin}} = \alpha_1 I'_{hkl} + \alpha_2 I''_{hkl} + \alpha_3 I'''_{hkl} + \dots = \sum_{i=1}^n \alpha_n I_{hkl}^n \quad (3.11)$$

donde  $I_{hkl}^{\text{twin}}$  es la intensidad de una reflexión determinada en un cristal maclado,  $n$  cada dominio distinto que forma parte del cristal y  $\alpha_n$  la contribución de cada dominio al cristal o fracción de macla.

Las diferentes orientaciones dentro de un cristal maclado se relacionan mediante una *operación de macla*, es decir, una operación de simetría que lleva un dominio a otro diferente. Un dominio se relaciona con los demás mediante varias operaciones de simetría, dando lugar a lo que se denomina *ley de macla*. Así pues, en resumen, un cristal maclado consta de dos o más dominios con diferente orientación que se relacionan entre sí mediante las leyes de macla (una ley por dominio).

El porcentaje de cada dominio dentro del cristal maclado puede variar, por lo que, además de por la ley de macla, cada dominio se caracteriza por su *fracción de macla* (parámetro  $\alpha$  en la Ecuación 3.11). Al dominio mayoritario se le suele considerar como el dominio original, calculándose las leyes de macla de los demás dominios respecto a este. Si todos los dominios presentan las mismas fracciones de macla o muy parecidas (es decir, no hay una orientación mayoritaria) se habla de *macla perfecta* y su resolución es, en la mayoría de los casos, imposible.

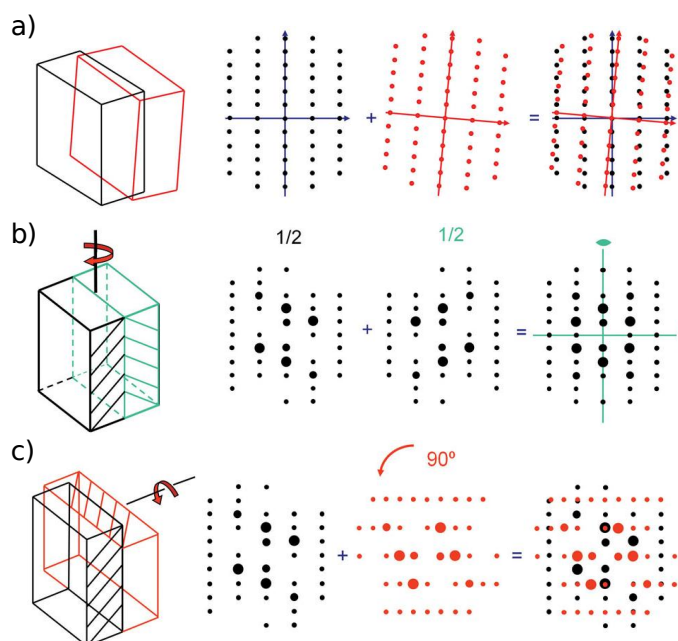
Existen diferentes clasificaciones de maclas, en base a si el interés es cristalográfico o mineralógico. A continuación se muestra la clasificación de los diferentes tipos de macla en función de las modificaciones que se producen a nivel de empaquetamiento y, por tanto, de patrón de difracción (Rupp, 2009):

- **Macla merohédrica:** Los diferentes dominios que componen el cristal maclado mantienen exactamente la misma red cristalina, por lo que el patrón de difracción resultante no mostrará defectos evidentes, sino reflexiones perfectamente ordenadas. No obstante, la simetría final será superior a la simetría real de cada orientación (Figura 3.27b). No se puede detectar mediante inspección visual, sino que se deben analizar las distribuciones de intensidades a lo largo de todo el patrón.
- **Macla pseudo-merohédrica:** Los diferentes dominios mantienen una red cristalina similar pero sin ser exactamente la misma. Las operaciones de macla equivalen a elementos de pseudo-simetría. Una estructura posee pseudo-simetría cuando una fracción significativa de los átomos que la componen satisfacen una simetría superior a la de la estructura completa (Giacovazzo *et al.*, 2011). Un ejemplo es la pseudo-simetría traslacional, que implica que una fracción significativa de los átomos se repita por una traslación que no pertenece al grupo espacial. Tanto los elementos de pseudo-simetría como la NCS pueden enmascarar un caso de macla, por lo que se complica la caracterización de la simetría real. Además, las pruebas que se realizan para analizar la presencia de macla pueden indicar



maclas negativas. Si bien este resultado es artificial, sirve como claro indicador de que el cristal presenta macla pseudo-merohédrica.

- **Macla no merohédrica:** A diferencia de los casos anteriores, donde las reflexiones se solapan dando un patrón de difracción de mayor simetría, una macla no merohédrica genera un patrón que puede ser separado y caracterizado para cada dominio (Figura 3.27c). Este tipo de macla surge al apilar orientaciones que crecen a partir de diferentes núcleos. No obstante, dependiendo del grado de solapamiento, algunos casos de macla no merohédrica se pueden tratar como merohédricos.



**Figura 3.27: Tipos de macla en función del patrón de difracción.** **a)** Una macla se caracteriza por la presencia de dos o más dominios con orientaciones diferentes, hecho que se refleja en una superposición de patrones. **b)** Macla merohédrica perfecta, con fracción de macla de 0,5. Los patrones se solapan en tal grado que su separación es inviable. Además, se han generado operaciones de simetría adicionales (líneas verdes) que enmascaran la simetría real de cada dominio individual. **c)** Macla no merohédrica, mostrando que la superposición de patrones no es tan grave como para que no puedan ser separados y analizados individualmente. Figura adaptada de Rupp, 2009.

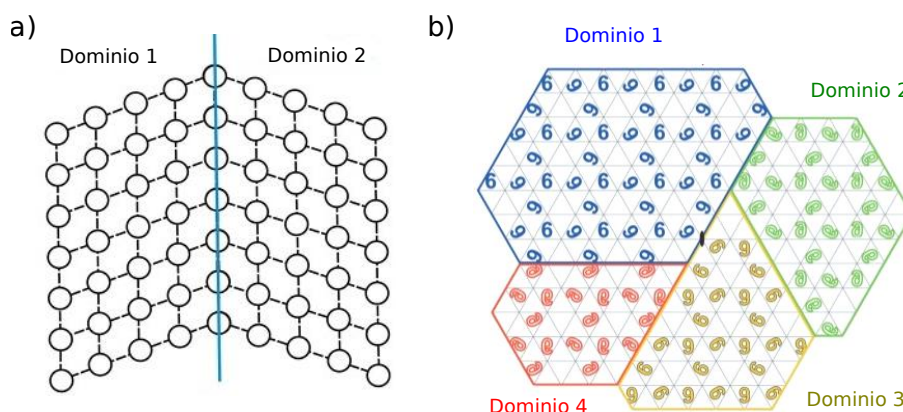
Además, se pueden establecer otros tipos de macla atendiendo a la cantidad de dominios que conforman el cristal. A continuación se describen dos tipos de macla de interés:

- **Macla hemihédrica:** Implica la presencia de un único dominio a mayores del dominio original, por lo que únicamente se requiere una ley de macla que describa el sistema (Figura 3.28a). Es el tipo más frecuente de macla merohédrica. Su presencia varía desde ser prácticamente inapreciable hasta presentar una fracción de macla de 0,5, es decir, macla hemihédrica perfecta.
- **Macla tetartohédrica:** Implica la presencia de cuatro dominios en el cristal maclado, por lo que se requiere un total de tres leyes de macla (Figura 3.28b). En el momento de publicación de este trabajo, apenas se han depositado una decena de estructuras en el PDB

que se hayan resuelto en presencia de macla tetartohédrica, lo que demuestra tanto la baja frecuencia de aparición de estos casos como el poco éxito en su detección y resolución.

La macla tetartohédrica es realmente un caso raro de macla merohédrica. Ni siquiera es aún claro que sea una macla exclusivamente merohédrica, ya que la mayoría de las veces que se ha detectado aparece acompañada de elementos de pseudo-simetría, lo que clasificaría la macla como pseudo-merohédrica (Roversi *et al.*, 2012). En el caso de las macromoléculas, únicamente los grupos espaciales  $P3$ ,  $P3_1$  y  $P3_2$  pueden albergar una macla tetartohédrica (Yeates & Yu, 2008), aunque otros grupos pueden, en casos excepcionales, permitir maclas pseudo-tetartohédricas (Anand *et al.*, 2007). Además, un cristal tetartohédrico puede ser analizado como un cristal maclado hemihédricamente en múltiples formas, de manera que un cristal tetartohédrico con grupo espacial  $P3$  real puede generar datos de difracción que se acerquen a los grupos  $P321$ ,  $P312$  o  $P6$  (Yeates & Yu, 2008). En general, la resolución de este tipo de maclas es bastante compleja. Mención especial requiere un caso resuelto recientemente en el que, además de tratarse de una macla tetartohédrica casi perfecta, la presencia de NCS traslacional de orden 7 obligó a construir 56 moléculas en el grupo espacial  $P1$  (Sliwiak *et al.*, 2014).

Desde el punto de vista del procesamiento de los datos de difracción, la macla provoca que los factores de estructura no sean tan precisos como en un caso normal, debido a su división en componentes de diferente peso. No obstante, si se obtienen fases razonables se puede llegar a un modelo inicial que permita retrasar hasta el proceso de refinamiento las correcciones sobre las intensidades observadas (Yeates & Yu, 2008). Como se ha visto anteriormente, el faseado experimental de un cristal normal (sin macla) permite obtener las fases reales de la proteína que contiene. En cambio, en el caso de cristales maclados, si la fracción de macla es muy elevada (o incluso perfecta), es probable que un modelo de reemplazo molecular sea la única opción que permita resolver la estructura, ya que las fases experimentales presentarán gran ambigüedad (Chandra *et al.*, 1999). No obstante, existen excepciones de casos extremos donde un faseado experimental ha permitido la resolución de datos altamente maclados (Dauter, 2003; Rudolph *et al.*, 2003; Terwisscha van Scheltinga *et al.*, 2003; Hillig & Renault, 2006).



**Figura 3.28: Tipos de macla en función del número de dominios.** a) Macla hemihédrica, compuesta por dos dominios, previsiblemente perfecta (fracción de macla = 0,5). b) Macla tetartohédrica, compuesta por cuatro dominios con diferente fracción de macla. El dominio 1 es el mayoritario. Figura adaptada de Yeates & Yu, 2008.

En la práctica, la resolución de datos maclados puede seguir dos estrategias diferentes. Por un lado, se puede intentar someter al conjunto de datos a un *desmaclado* que permita separar las diferentes reflexiones y contribuciones en diferentes conjuntos de datos, para posteriormente proceder a realizar la resolución estructural como si de datos de un sólo dominio se tratara. No obstante, el desmaclado no siempre es efectivo y menos aún en casos cercanos a maclas perfectas (Dauter, 2003). Por otro lado, la estrategia más habitual consiste en refinar las estructuras indicando tanto la fracción como la ley de macla de cada dominio, de tal forma que los propios programas de refinamiento puedan discriminar unas reflexiones de otras y calcular las contribuciones que cada dominio ejerce sobre la estructura final.

Se puede sospechar de la existencia de maclas si la celdilla unidad es demasiado pequeña y no permite que la molécula quepa en ella. También es probable que exista una macla si los datos son buenos, la estructura se resuelve con buenos mapas pero los factores  $R$  de refinamiento se estabilizan en valores elevados. Como es habitual en cada etapa del proceso de resolución estructural, existen diferentes indicadores capaces de determinar la presencia, el tipo y el grado de macla de un cristal. Algunos se basan en la totalidad de las intensidades, pudiéndose aplicar tanto en la baja como en la alta simetría, pero los más efectivos se basan en la comparación entre intensidades directamente relacionadas por la macla, teniendo que ser utilizados sobre datos procesados en su grupo espacial real, de menor simetría. Un análisis exhaustivo de las diferentes pruebas se puede encontrar en Dauter, 2003. En esta sección nos limitaremos a mostrar los valores frontera de cada estadístico, enumerados en la Tabla 3.3.

**Tabla 3.3:** Indicadores estadísticos para el análisis de maclas.

ESTADÍSTICO	SIN MACLA	MACLA PERFECTA
$\langle I^2 \rangle / \langle I \rangle^2$	2,000	1,500
$\langle F \rangle^2 / \langle F^2 \rangle$	0,785	0,885
$\langle  E^2 - 1  \rangle$	0,736	0,541
$\langle  L  \rangle$	0,500	0,375
$\langle L^2 \rangle$	0,333	0,200

Además, algunas pruebas permiten calcular la fracción de macla ( $\alpha$ ) para cada dominio por separado, algo especialmente útil en el caso de macla tetartohédrica donde es fundamental conocer la contribución de cada dominio de forma individual. Estas pruebas son el test  $H$  (Yeates, 1988; Yeates, 1997) y el test de Britton (Britton, 1972).

Por último, es importante tener en cuenta que cada dominio se comporta como un cristal independiente. Esto implica que, al igual que sucede con monocristales, la calidad de la difracción de cada dominio puede variar, obteniendo diferentes resoluciones para cada dominio. En la práctica, este hecho se traduce en un salto en la calidad de la difracción significativo a partir de determinada capa de resolución. Así pues, unos datos maclados se pueden procesar con diferentes cortes de resolución, obteniendo diferentes contribuciones (fracciones  $\alpha$ ) de los dominios implicados. Este hecho es de especial interés en cristales merohédricos hemihédricos

perfectos, donde se puede intentar rebajar la contribución de algún dominio para conseguir una macla parcial resoluble.

El principal programa utilizado a lo largo de este trabajo para analizar la presencia y tipo de maclas fue `phenix.xtriage`, perteneciente al paquete PHENIX (Adams *et al.*, 2010). Los intentos de desmaclado de datos se realizaron con `Detwin`, perteneciente al paquete CCP4 (Winn *et al.*, 2011).

---

---

## RESULTADOS Y DISCUSIÓN

---

---



---

# DETERMINACIÓN ESTRUCTURAL Y FUNCIONAL DE MUR A1

---

## Índice

4.1	Caracterización de MurA1 . . . . .	74
4.2	Obtención de MurA1 <sub>D39</sub> y MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	74
4.3	Cristalización de MurA1 <sub>D39</sub> . . . . .	75
4.4	Determinación estructural de MurA1 <sub>D39</sub> . . . . .	76
4.5	Análisis estructural de MurA1 <sub>D39</sub> . . . . .	78
4.5.1	Dominios estructurales . . . . .	78
4.5.2	Empaquetamiento cristalino . . . . .	79
4.5.3	Agitación térmica . . . . .	80
4.5.4	Solvente . . . . .	80
4.6	Cristalización de MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	80
4.7	Determinación estructural de MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	81
4.8	Análisis estructural de MurA1 <sub>Hungary<sup>19A</sup></sub> . . . . .	83
4.8.1	Estructura general . . . . .	83
4.8.2	Empaquetamiento cristalino . . . . .	84
4.8.3	Agitación térmica . . . . .	84
4.8.4	Otras moléculas y solvente . . . . .	85
4.9	Cristalización de MurA1 <sub>D39</sub> y MurA1 <sub>Hungary<sup>19A</sup></sub> en presencia de ligandos . . .	86
4.10	Relación estructura-función . . . . .	87
4.10.1	Análisis comparado de estructuras . . . . .	87
4.10.2	Reconocimiento de sustratos y sitio activo . . . . .	89
4.10.3	Inhibición mediante fosfomicina y citrato . . . . .	91
4.10.4	Hipótesis sobre el mecanismo de reacción . . . . .	94
4.10.5	MurA1 <sub>Hungary<sup>19A</sup></sub> y heterorresistencia a fosfomicina . . . . .	95

---

## 4.1. Caracterización de MurA1

Las bacterias Gram positivas contienen dos genes (*murA1* y *murA2*) que codifican para dos isoformas de la proteína MurA (Du *et al.*, 2000). La proteína MurA1 de *S. pneumoniae* consta de 419 aminoácidos y un peso molecular teórico de 45,0 kDa y un pI teórico de 5,2. En base a su similitud e identidad con la proteína MurA de *Enterobacter cloacae* (59,1 y 43,9% respectivamente), se espera que MurA1 esté formada por dos dominios globulares unidos entre sí, ubicándose el sitio catalítico entre ambos. Además se prevé la existencia de un lazo flexible que contenga la Cys116 catalítica (Figura 4.1).

10	20	30	40	50	60
MRKIVINGGL	PLQGEITISG	AKNSVVALIP	AAILADDVVT	LDCVPDISDV	ASLVEIMELM
70	80	90	100	110	120
GATVKRYDDV	LEIDPRGVQN	IPMPYBKINS	LRASYFYFGS	LLGRFGEATV	GLPGGC <sup>■</sup> DLGP
130	140	150	160	170	180
RPIDLHLKAF	EAMGATASYE	GDNMKLSAKD	TGLHGASIYM	DTVSVGATIN	TMIAAVKANG
190	200	210	220	230	240
RTIIENAARE	PEIIDVATLL	NNMGAHIRGA	GTNIIIDGV	ERLHGTRHQV	IPDRIEAGTY
250	260	270	280	290	300
ISLAAAVGKG	IRINNVLYEH	LEGFIAKLEE	MGVRMTVSED	SIFVEEQSNL	KAINIKTAPY
310	320	330	340	350	360
PGFATDLQQP	LTPLLLRANG	RGTIVDTIYE	KRVNHVFELA	KMDADISTTN	GHILYTGGRD
370	380	390	400	410	
LRGASVKATD	LRAGAALVIA	GLMAEGKTEI	TNIEFILRGY	SDIIEKLRLN	GADIRLVED

**Figura 4.1:** Secuencia de la proteína MurA1 de *S. pneumoniae* D39. Los dos dominios teóricos se han coloreado en naranja y azul. La cisteína catalítica se ha destacado con un recuadro negro.

La única diferencia a nivel de secuencia entre las proteínas MurA1 de las cepas D39 y Hungary<sup>19A</sup> de *S. pneumoniae* es la sustitución Ala364Thr. A continuación se describen los experimentos llevados a cabo y los resultados obtenidos durante la determinación estructural de la proteína MurA1 en las cepas D39 y Hungary<sup>19A</sup> de *S. pneumoniae*, designadas como MurA1<sub>D39</sub> y MurA1<sub>Hungary<sup>19A</sup></sub> respectivamente.

## 4.2. Obtención de MurA1<sub>D39</sub> y MurA1<sub>Hungary<sup>19A</sup></sub>

Las proteínas MurA1<sub>D39</sub> y MurA1<sub>Hungary<sup>19A</sup></sub> fueron clonadas, expresadas y purificadas por el grupo de la Prof. Kathrin Mühlemann y la Dra. Lucy J. Hathaway (Instituto de Enfermedades Infecciosas, Facultad de Medicina, Universidad de Berna, Suiza). El gen *murA1* de las cepas D39 y Hungary<sup>19A</sup> se clonó separadamente utilizando el plásmido pGEX-6p-1, insertándose entre los sitios *Xho*I y *Bam*HI. La defosforilación se realizó con la fosfatasa SAP y la ligación con una ligasa de DNA del fago T4 a 4 °C durante 3 días. La transformación se realizó en células *E. coli* SURE y la selección de los transformantes en medio LBA con 100 µg/mL de ampicilina.



Posteriormente, los plásmidos se aislaron mediante el sistema Promega Wizard y se utilizaron para transformar células *E. coli* BL21 competentes. La secuencia clonada de *murA1* se confirmó mediante secuenciación.

Las células *E. coli* BL21 que contenían el gen de interés se crecieron en medio LB líquido con 100  $\mu\text{g/mL}$  de ampicilina hasta alcanzar una  $\text{DO}_{600}$  de 0,5 a 37 °C. A continuación se produjo la inducción de la expresión mediante la adición de 0,1 mM de isopropil- $\beta$ -D-1-tiogalactopiranosido (IPTG), manteniendo el cultivo a 20 °C durante 16 horas.

Tras la expresión, se recogieron las células mediante centrifugación a 4000 g durante 10 minutos y se lisaron en hielo con 25 mL de un tampón de lisis consistente en 50 mM  $\text{NaH}_2\text{PO}_4$  pH 8,0, 300 mM NaCl, 10 mM imidazol y 1 mg/mL lisozima (Roche, Suiza) durante 30 minutos. Las proteasas se inhibieron mediante la adición de 1  $\mu\text{g/mL}$  de aprotinina, 1  $\mu\text{g/mL}$  de leupeptina, 1  $\mu\text{g/mL}$  de pepstatina y 1 mM de fluoruro de fenilmetilsulfonilo (PMSF).

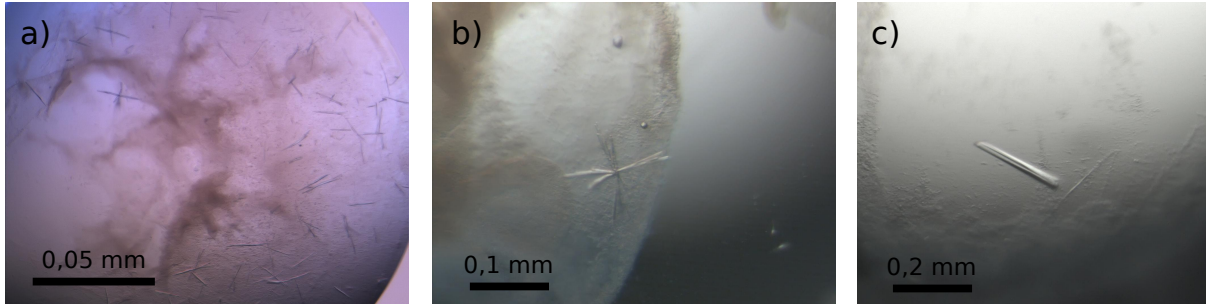
La purificación de cada proteína se realizó mediante una cromatografía en la columna Glutathione-Sepharose® 4 Fast Flow (GE Healthcare, Suecia), utilizando la PreScission Protease (GE Healthcare, Suecia) para eliminar la etiqueta de purificación GST. Para separar la proteína de las esferas se utilizaron columnas Qiagen Polypropylene. Finalmente, se intercambió el tampón de las proteínas por 10 mM Tris-HCl pH 7,5 en filtros Amicon Ultra-4 (Millipore, Irlanda). La pureza y la identidad de cada proteína extraída se confirmó mediante SDS-PAGE y huella peptídica respectivamente.

### 4.3. Cristalización de MurA1<sub>D39</sub>

La cristalización de la proteína MurA1<sub>D39</sub> se inició con la búsqueda de condiciones iniciales mediante técnicas de alto rendimiento, descritas en la Sección 3.2.5. Los paquetes comerciales ensayados durante esta etapa fueron Crystal Screen, Crystal Screen 2, INDEX, JCSG, PACT Suite y JBScreen Classic 1-4 (Tabla 3.1). En cada ensayo se utilizaron 250 nL de proteína y 250 nL de cada precipitante, a una temperatura de 18 °C. Se ensayaron tres concentraciones distintas de proteína (5,4 mg/mL, 7,2 mg/mL y 8,3 mg/mL) en 10 mM Tris-HCl pH 7,0, con la intención de ampliar el rango de condiciones de cristalización por variación de la concentración.

En esta primera etapa se observaron cristales en una condición que contenía 200 mM  $\text{CaCl}_2$ , 100 mM HEPES pH 7,5 y 30 % (p/v) PEG 4000, encontrándose la proteína a 5,4 mg/mL. Dichos cristales presentaban una morfología de aguja fina de pequeño tamaño, lo que imposibilitaba su utilización en un experimento de difracción de rayos X (Figura 4.2a). Así pues, se realizó una optimización de la cristalización mediante difusión de vapor en gota colgante ensayando gradientes de la condición original, variando la concentración de uno de los componentes en cada nueva condición y probando diferentes puntos de pH en torno al pI de la proteína. Además, en algunos experimentos se sustituyó el PEG 4000 por PEG 400, buscando modificar la dinámica de agregación de la proteína. Finalmente, se obtuvieron cristales individuales de mayor tamaño, a 200 mM  $\text{CaCl}_2$ , 100 mM HEPES pH 7.5 y 24 % (p/v) PEG 4000 (Figura 4.2b). En una optimización posterior se consiguieron los mejores cristales de MurA1<sub>D39</sub>, a una concentración de 14 mg/mL y frente a 50 mM  $\text{CaCl}_2$ , 100 mM HEPES pH 7.5 y 24 % (p/v) PEG 4000 (Figura

4.2c). Por último, se llevó a cabo la cristalización de MurA1<sub>D39</sub> en presencia de aditivos y detergentes, pero no se observaron mejoras significativas en la calidad de los cristales.



**Figura 4.2:** Cristales de MurA1<sub>D39</sub>. a) Aguja pequeñas en una gota de 500 nL. b) Cristales de mayor tamaño, en una gota de 2  $\mu$ L optimizada. c) Cristales finales que dieron lugar a la resolución de la estructura.

#### 4.4. Determinación estructural de MurA1<sub>D39</sub>

Los diferentes cristales obtenidos a lo largo de las etapas de optimización se crioprotegieron en 20 % (v/v) glicerol y se llevaron a las líneas ID14-1 e ID14-4 del ESRF y a la línea X06DA (PXIII) del SLS, donde finalmente se obtuvo un patrón de calidad suficiente para la resolución de la estructura a nivel atómico (Tabla 4.1).

**Tabla 4.1:** Datos y estadísticas de difracción de MurA1<sub>D39</sub>.

PARÁMETRO	MUR A1 <sub>D39</sub>
Grupo espacial	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
$a, b, c$ (Å)	46,62, 81,11, 104,88
$\alpha = \beta = \gamma$ (°)	90,0, 90,0, 90,0
Resolución (Å)	44,0 – 2,9 (3,06 – 2,9)
Longitud de onda (Å)	1,000
Línea	PXIII (SLS)
Temperatura (K)	100
Nº de reflexiones	
Totales	185839
Únicas	14373
$\langle I/\sigma I \rangle$	7,0 (1,9)
$R_{\text{pim}}$	0,114 (0,427)
Compleitud (%)	100 (99,7)
Multiplicidad	6,5 (6,2)

Los valores entre paréntesis hacen referencia a la última capa de resolución.

Los diferentes patrones de difracción obtenidos a lo largo de los ensayos de optimización se indexaron, integraron y escalaron con los programas XDS, iMosflm y Scala. Los mejores cristales de la MurA1<sub>D39</sub> difrataron a 2,9 Å, perteneciendo al grupo espacial P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> con celdilla  $a = 46,62$  Å,  $b = 81,11$  Å,  $c = 104,88$  Å,  $\alpha = \beta = \gamma = 90^\circ$ . El coeficiente de Matthews

resultante fue 2,18 Å<sup>3</sup>/Da, indicando la presencia de un único monómero en la unidad asimétrica y un porcentaje de solvente del 43,5 %.

Debido a la existencia previa de la estructura tridimensional de la proteína MurA de *E. cloacae* (código PDB 1NAW; Schönbrunn *et al.*, 1996), el método de faseado elegido fue el reemplazo molecular, realizado con el programa Phaser. La similitud de secuencia entre la proteína MurA1<sub>D39</sub> y la proteína MurA de *E. cloacae* es del 59,4 %, mientras que su identidad es del 44,2 %. Estos valores y el hecho de que ambas proteínas realicen la misma función pero en microorganismos distintos son razones más que suficientes para abordar el faseado mediante reemplazo molecular.

Ya que la proteína MurA de *E. cloacae* consta de dos dominios globulares unidos por una región relativamente flexible, era de esperar que MurA1<sub>D39</sub> presentara un plegamiento similar. Así pues, el reemplazo molecular se realizó aportando los dos dominios de la MurA de *E. cloacae* por separado (residuos Ala21 – Leu229 y Arg232 – Glu419). De esta forma se evita restringir la flexibilidad de la proteína MurA1<sub>D39</sub> a los valores de la MurA modelo. El resultado obtenido permitió determinar con exactitud la geometría y el plegamiento de la cadena principal de la MurA1<sub>D39</sub>, mientras que las cadenas laterales fueron corregidas con el programa Chainsaw.

La solución obtenida tras el reemplazo molecular y la corrección de las cadenas laterales se refinó con el programa phenix.refine. En primer lugar se realizó un ciclo de refinamiento de cuerpo rígido para ajustar el modelo general a la densidad, seguido de 10 ciclos de enfriamiento simulado que ajustaron la geometría de la mayoría de las cadenas laterales generadas con Chainsaw. Posteriormente, se corrigieron de forma manual, con el programa Coot, zonas críticas de la estructura como lazos flexibles expuestos al solvente y cadenas laterales fuera de densidad, seguido de varios ciclos adicionales de refinamiento de coordenadas y de factores de agitación térmica. Los factores *R* de refinamiento finales fueron *R*<sub>work</sub>= 0,18 / *R*<sub>free</sub>= 0,24. La Tabla 4.2 contiene las estadísticas relativas a la etapa de refinamiento de MurA1<sub>D39</sub>.

**Tabla 4.2:** Estadísticas de refinamiento de MurA1<sub>D39</sub>.

PARÁMETRO	MUR A1 <sub>D39</sub>
Resolución (Å)	44,0 – 2,9
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0,18/0,24
Nº de átomos	
Proteína (no H)	3118
Agua	14
Factores B (Å <sup>2</sup> )	53,8
RMSD de enlaces	
Distancias (Å)	0,002
Ángulos (°)	0,504
Ramachandran (%)	
Favorables	97,1
Desfavorables	0,5
<b>Código PDB</b>	<b>3ZH3</b>

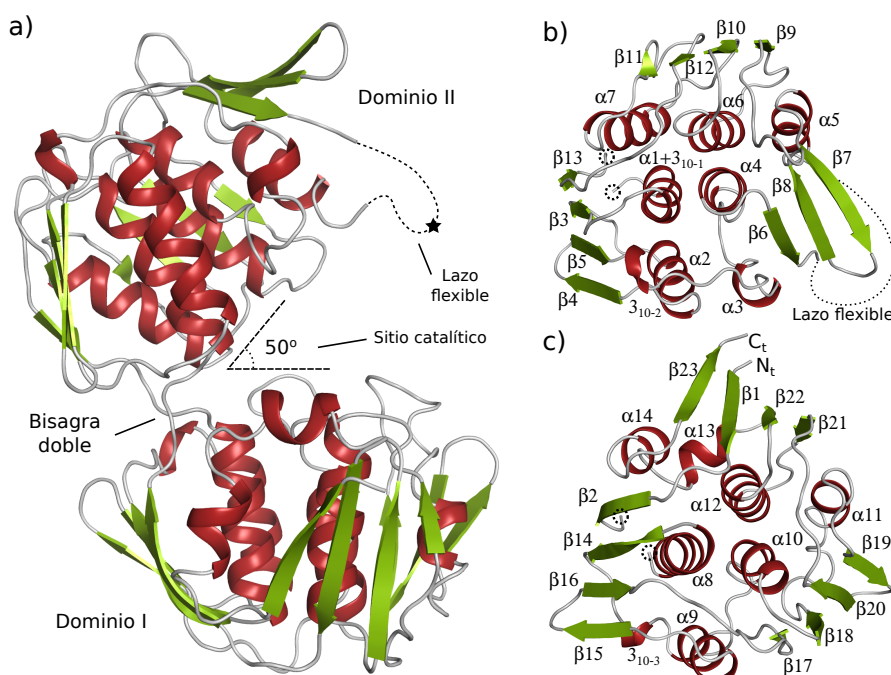
Los valores correspondientes a la RMSD de enlaces y a los requerimientos de Ramachandran indican una geometría óptima para el modelo construido, mientras que los factores B generales muestran una buena correlación con los valores teóricos esperados para la resolución a la que se han recogido los datos de difracción.

## 4.5. Análisis estructural de MurA1<sub>D39</sub>

La calidad de los mapas obtenidos permitió construir la estructura de MurA1<sub>D39</sub> desde la metionina inicial hasta el Asp419 final, salvo una zona intermedia correspondiente con un lazo flexible (de Gly114 a Pro120).

### 4.5.1. Dominios estructurales

La proteína MurA1<sub>D39</sub> de *S. pneumoniae* está constituida por dos dominios globulares (dominios I y II) unidos por una bisagra doble (residuos Ser – Gly e Ile – Pro – Asp en las posiciones 19 y 231 respectivamente), siendo su separación angular de aproximadamente 50° (Figura 4.3a). Cada dominio está formado por un total de 8 hélices de tipo  $\alpha$  y  $3_{10}$  rodeadas por 3 láminas  $\beta$  de 4 de hebras  $\beta$  cada una (a excepción de la lámina  $\beta_6 - \beta_7 - \beta_8$ , que contiene el lazo flexible). Los dominios ocupan aproximadamente el volumen de una esfera de 20 Å de radio. El dominio II está formado por los residuos Ala21 a Val230 e incluye el lazo flexible que no ha podido ser construido (Figura 4.3c), mientras que el dominio I contiene el extremo N-terminal

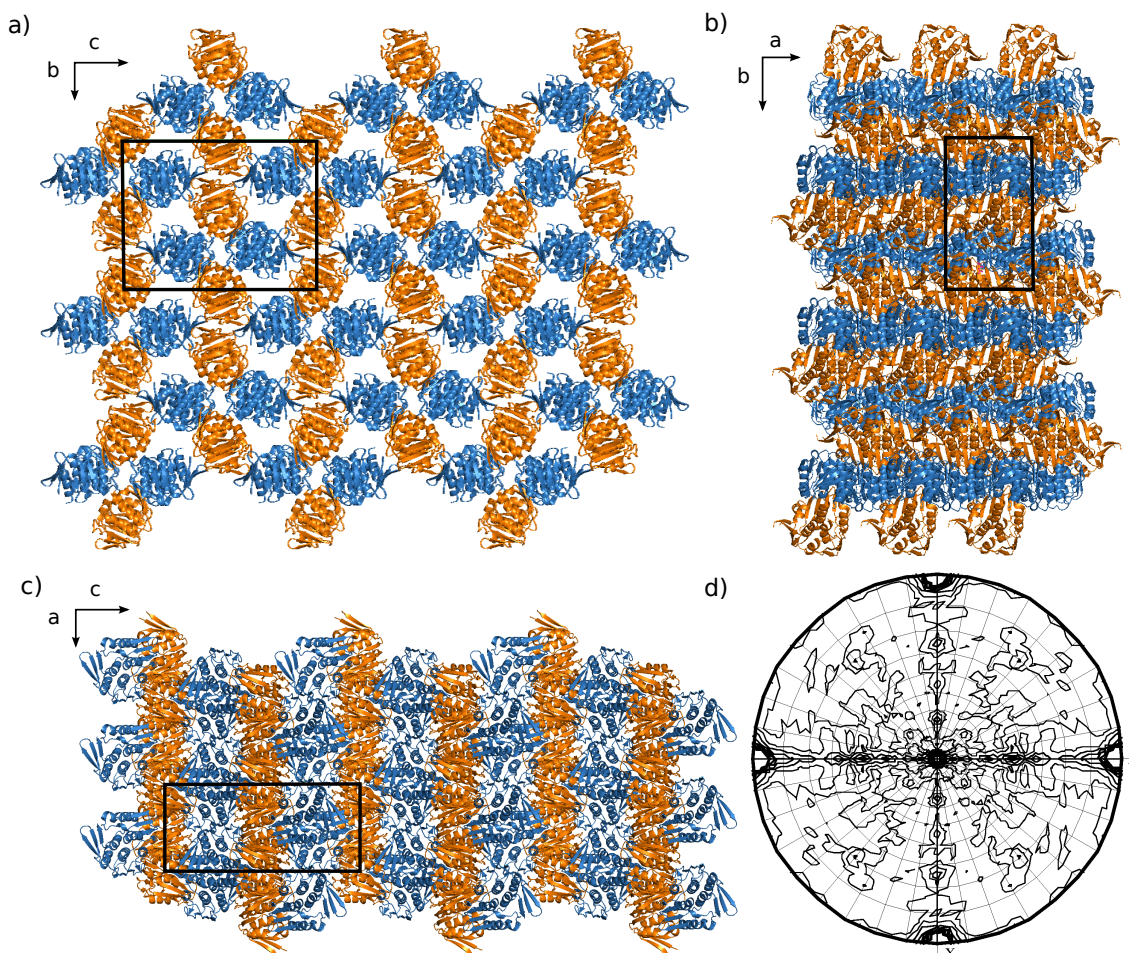


**Figura 4.3: Estructura general de MurA1<sub>D39</sub>.** a) Estructura global de MurA1<sub>D39</sub> mostrando los dominios que la componen, su separación y la localización tanto de la bisagra doble como del lazo flexible. La estrella hace referencia a la localización teórica de la Cys116 catalítica. b) Estructura global del dominio II, mostrando la ubicación de sus elementos de estructura secundaria y la localización teórica del lazo flexible. c) Estructura global del dominio I, mostrando los extremos N y C-terminales ( $N_t$  y  $C_t$  respectivamente). Los círculos punteados de los paneles b y c muestran los residuos frontera de cada dominio a la altura de la bisagra doble.

(residuos Met1 a Ile18) y la región C-terminal correspondiente a los residuos Arg234 a Asp419 (Figura 4.3c). El sitio catalítico se sitúa entre los dos dominios, tomando residuos de ambas partes. La Cys116, presuntamente implicada en el mecanismo catalítico, se encuentra en el lazo flexible.

#### 4.5.2. Empaquetamiento cristalino

La proteína MurA1<sub>D39</sub> ha cristalizado en el grupo espacial ortorrómbico  $P2_12_12_1$ , lo que implica la existencia de 3 ejes binarios helicoidales. Al aplicar las operaciones de simetría se determina que hay un total de 4 moléculas en la celdilla unidad. Como se puede observar en la Figura 4.4, la disposición de las diferentes moléculas en el cristal es tal que los dominios I de unas moléculas interaccionan con los dominios II de las vecinas y viceversa, creando un empaquetamiento estable. Además, en la vista perpendicular al plano  $(b, c)$  (Figura 4.4a) se observan unos canales de solvente, coincidentes con la región del lazo flexible con la Cys116. Este hecho justifica su flexibilidad y el hecho de que no se haya obtenido una densidad electrónica fiable que permita su construcción.

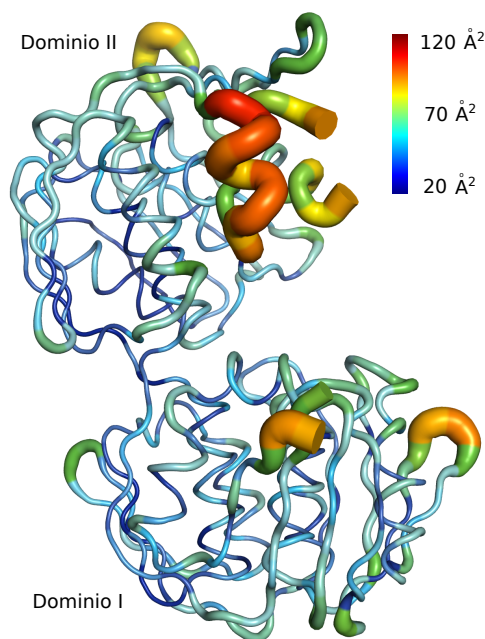


**Figura 4.4: Simetría y empaquetamiento de MurA1<sub>D39</sub>.** El dominio I se representa en naranja, mientras que en azul se representa el dominio II. El recuadro negro hace referencia a la celdilla unidad, indicándose además la orientación de cada eje en cada caso. **a)** Plano  $(b, c)$ . **b)** Plano  $(a, b)$ . **c)** Plano  $(a, c)$ . **d)** Función de autorrotación calculada para  $\chi = 180^\circ$  mediante Molrep, mostrando los picos característicos de un grupo de Laue  $mmm$ .



### 4.5.3. Agitación térmica

Los factores B medios globales, situados en  $53,8 \text{ \AA}^2$ , indican una estabilidad aceptable a lo largo de toda la estructura. En concreto, los factores B medios de la cadena principal se sitúan en torno a  $51,5 \text{ \AA}^2$ , mientras que los factores B medios correspondientes a las cadenas laterales rondan los  $56,3 \text{ \AA}^2$ . No obstante, la región Pro84 – Arg92 presenta factores B en torno a  $120 \text{ \AA}^2$ , lo que es reflejo de una gran agitación e inestabilidad térmica (Figura 4.5). Además, esta región se ubica en la vecindad del lazo flexible que contiene la Cys116 catalítica, ausente en el modelo (Pro113 a Arg121).



**Figura 4.5: Agitación térmica de MurA1<sub>D39</sub>.** Representación gráfica de la agitación térmica de MurA1<sub>D39</sub> en base a los factores B medios de cada residuo. Se puede observar una mayor agitación en la región de la hélice  $\alpha_3$  (Pro84 – Arg92, de 100 a  $120 \text{ \AA}^2$ ) y en la correspondiente al lazo flexible (siendo únicamente visibles los extremos, con factores B cercanos a  $100 \text{ \AA}^2$ ). Otras zonas flexibles menos significativas se corresponden con lazos cortos expuestos al solvente (entre 70 y  $90 \text{ \AA}^2$ ).

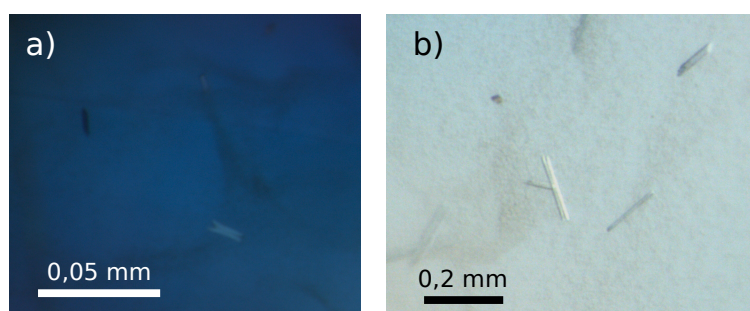
### 4.5.4. Solvente

En lo que respecta a la modelización del solvente, la relativamente baja resolución de los datos ( $2,9 \text{ \AA}$ ) no permitió construir más de 14 moléculas de agua, situadas en las regiones más estables de la proteína o en zonas de contacto con moléculas vecinas dentro del cristal. No se encontró ninguna densidad relacionada con otras moléculas presentes en la condición de cristalización.

## 4.6. Cristalización de MurA1<sub>Hungary</sub><sup>19A</sup>

El procedimiento de cristalización de MurA1<sub>Hungary</sub><sup>19A</sup> comenzó con la búsqueda de unas condiciones iniciales adecuadas para la formación de cristales. La presencia de una mutación puntual como elemento diferencial entre MurA1<sub>Hungary</sub><sup>19A</sup> y MurA1<sub>D39</sub> podía implicar un plegamiento diferente de alguna región de la proteína, lo que supondría diferencias en el

empaquetamiento y el proceso de cristalización. Por tanto, se procedió a realizar una búsqueda de condiciones iniciales mediante técnicas de alto rendimiento, utilizando los paquetes comerciales Crystal Screen, Crystal Screen 2, INDEX, JCSG, PACT Suite y JBScreen Classic 1-4 (Tabla 3.1). En cada ensayo se utilizaron 250 nL de proteína a 7,5 mg/mL y 250 nL de cada precipitante, a una temperatura de 18 °C. Se obtuvieron unos primeros cristales en 200 mM acetato amónico, 100 mM citrato sódico tribásico pH 5,6 y 30 % (p/v) PEG 4000 (Figura 4.6a). La optimización de los cristales se realizó mediante el método de difusión de vapor en gota sentada, ensayando gradientes de concentración de cada uno de los componentes de la condición de cristalización original y a diferentes puntos de pH en torno a su pI. Los mejores cristales se obtuvieron en 200 mM acetato amónico, 100 mM citrato sódico tribásico pH 5,6 y 35 % (p/v) PEG 4000 (Figura 4.6b).



**Figura 4.6: Cristales de MurA1<sub>Hungary19A</sub>.** a) Primeros cristales obtenidos en los ensayos iniciales de cristalización. b) Cristales de mayor tamaño obtenidos tras la etapa de optimización.

#### 4.7. Determinación estructural de MurA1<sub>Hungary19A</sub>

Los cristales obtenidos en las etapas anteriores se crioprotegieron en 20 % (v/v) glicerol y se midieron en la línea X06DA (PXIII) del SLS. Los datos fueron indexados, integrados y escalados con los programas XDS y Scala. En la Tabla 4.3 se muestran los datos y estadísticas de difracción del mejor cristal de MurA1<sub>Hungary19A</sub>.

El faseado se realizó mediante reemplazo molecular, utilizando como modelo la estructura de MurA1<sub>D39</sub> previamente resuelta. El coeficiente de Matthews para los datos de MurA1<sub>Hungary19A</sub> era 1,88 Å<sup>3</sup>/Da, lo que indicaba la presencia de un único monómero y un 34,8 % de solvente. Inicialmente no se encontró solución de reemplazo para la totalidad de la estructura, hecho que, junto con el cambio en el grupo espacial y los parámetros de celdilla unidad, indicaba algún cambio conformacional importante en MurA1<sub>Hungary19A</sub> respecto a MurA1<sub>D39</sub>. Así pues, se realizó un segundo reemplazo molecular utilizando los dominios separados de MurA1<sub>D39</sub> como modelos independientes, con el fin de evitar las variaciones producidas por la flexibilidad de la estructura. Con esta aproximación fue posible encontrar una solución óptima, utilizada como base para la construcción de la estructura final de MurA1<sub>Hungary19A</sub>.

Tras corregir manualmente la mutación puntual, se procedió a realizar un ciclo de refinamiento de cuerpo rígido, seguido de 10 ciclos de enfriamiento simulado, con la intención de generar posiciones y factores de agitación térmica adecuados para los datos experimentales

de MurA1<sub>Hungary19A</sub> y borrar el posible sesgo proveniente del modelo de MurA1<sub>D39</sub>. Finalmente, se sucedieron varios ciclos combinados de construcción manual y refinamiento de coordenadas atómicas y factores de ocupación térmica. Los factores  $R$  de refinamiento convergieron a  $R_{\text{work}} = 0,16$  /  $R_{\text{free}} = 0,20$ . En la Tabla 4.4 se muestran las estadísticas de refinamiento de MurA1<sub>Hungary19A</sub>. Los refinamientos se realizaron con el programa phenix.refine, mientras que la construcción manual se realizó con el programa Coot.

**Tabla 4.3:** Datos y estadísticas de difracción de MurA1<sub>Hungary19A</sub>.

PARÁMETRO	MUR A1 <sub>HUNGARY19A</sub>
Grupo espacial	P2 <sub>1</sub> 2 <sub>1</sub> 2
$a, b, c$ (Å)	106,58, 70,36, 45,24
$\alpha = \beta = \gamma$ (°)	90,0, 90,0, 90,0
Resolución (Å)	45,2 – 1,8 (1,9 – 1,8)
Longitud de onda (Å)	1,000
Línea	PXIII (SLS)
Temperatura (K)	100
Nº de reflexiones	
Totales	420364
Únicas	32258
$\langle I/\sigma I \rangle$	16,3 (2,3)
$R_{\text{pim}}$	0,039 (0,351)
Compleitud (%)	99,7 (99,7)
Multiplicidad	12,7 (13,0)

Los valores entre paréntesis hacen referencia a la última capa de resolución.

**Tabla 4.4:** Estadísticas de refinamiento de MurA1<sub>Hungary19A</sub>.

PARÁMETRO	MUR A1 <sub>HUNGARY19A</sub>
Resolución (Å)	45,2 – 1,8
$R_{\text{work}}/R_{\text{free}}$	0,16/0,20
Nº de átomos	
Proteína (no H)	3138
Citrato	18
PEG	17
Agua	348
Factores B (Å <sup>2</sup> )	
Proteína	20,8
Citrato	27,2
PEG	45,6
RMSD de enlaces	
Distancias (Å)	0,007
Ángulos (°)	1,07
Ramachandran (%)	
Favorables	97,3
Desfavorables	0,5
<b>Código PDB</b>	<b>3ZH4</b>



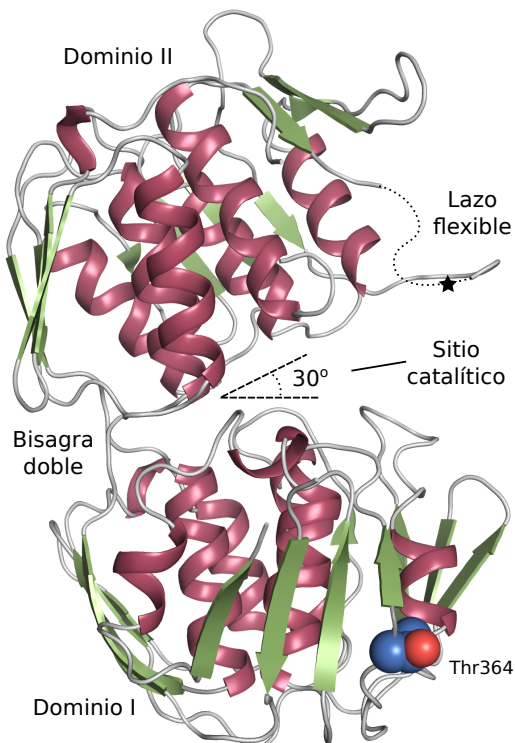
Tanto los factores B de la proteína como los RMSD de enlaces y los valores de Ramachandran se corresponden con los valores teóricos esperables para una estructura de la resolución obtenida para MurA1<sub>Hungary</sub><sup>19A</sup>.

## 4.8. Análisis estructural de MurA1<sub>Hungary</sub><sup>19A</sup>

Los mapas de densidad obtenidos para MurA1<sub>Hungary</sub><sup>19A</sup> permitieron construir la proteína desde la metionina inicial hasta el Glu418 (penúltimo residuo de la cadena), salvo dos regiones correspondientes a los residuos Tyr85 – Lys87 y Pro113 – Cys116.

### 4.8.1. Estructura general

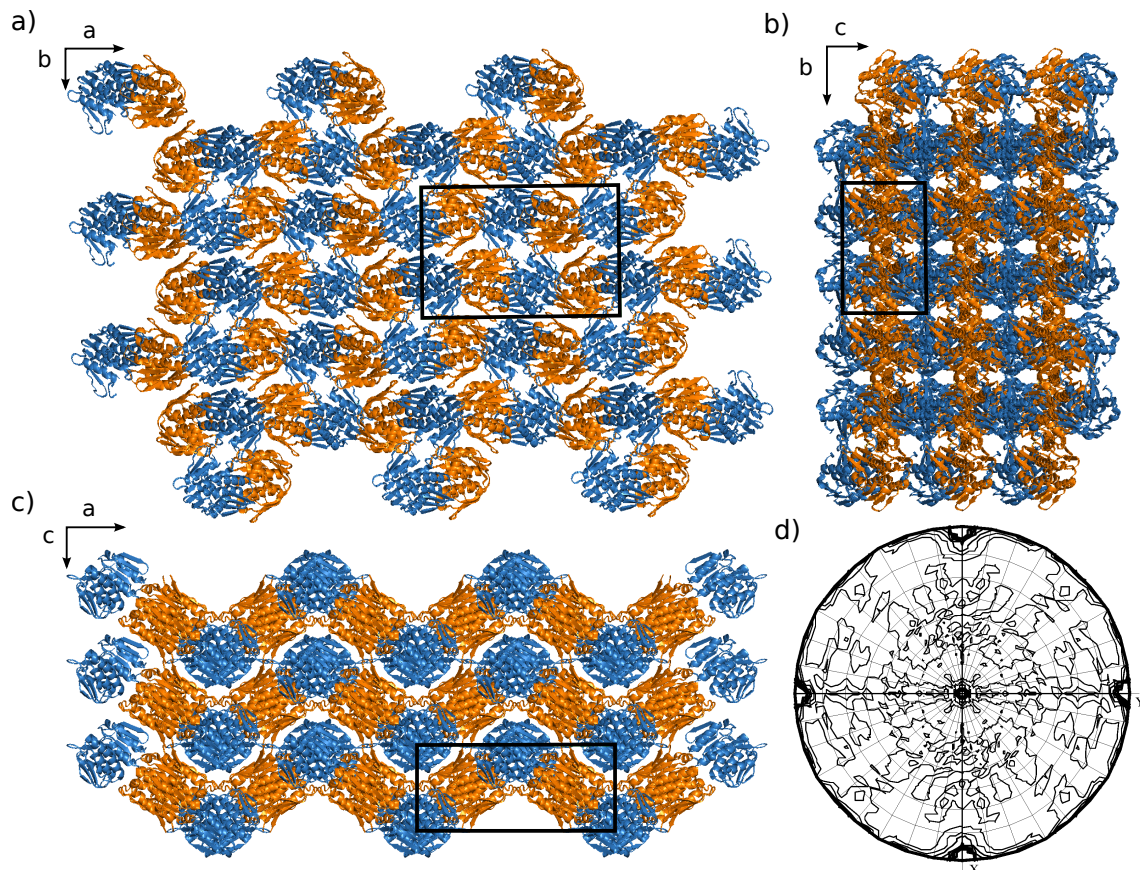
La proteína MurA1<sub>Hungary</sub><sup>19A</sup> de *S. pneumoniae* presenta la misma organización en dominios que la proteína MurA1<sub>D39</sub> previamente descrita (Sección 4.5.1). No obstante, si bien el plegamiento se mantiene invariante, la separación angular entre ambos dominios es sensiblemente menor que en el caso de MurA1<sub>D39</sub>, alcanzando un valor en torno a 30° (Figura 4.7). Este hecho explica la falta de éxito en los reemplazos moleculares en los que se utilizó la proteína MurA1<sub>D39</sub> completa como modelo. Pese a la mejor calidad de los datos con respecto a MurA1<sub>D39</sub>, el lazo flexible que contiene la Cys116 tampoco pudo ser modelado en su totalidad. La principal novedad de la estructura de MurA1<sub>Hungary</sub><sup>19A</sup> respecto a MurA1<sub>D39</sub> es la presencia de la mutación Ala<sub>364</sub>Thr (Figura 4.7), situada justo antes del inicio de la  $\beta_{21}$  del dominio I.



**Figura 4.7: Estructura general de MurA1<sub>Hungary</sub><sup>19A</sup>.** Estructura general de la proteína MurA1<sub>Hungary</sub><sup>19A</sup> mostrando los dominios que la componen, su separación y la localización tanto de la bisagra doble como del lazo flexible (la localización teórica de la Cys116 se indica con una estrella). En esferas se muestra la Thr364, única diferencia en composición respecto a MurA1<sub>D39</sub>, que tiene una alanina en esta posición.

### 4.8.2. Empaquetamiento cristalino

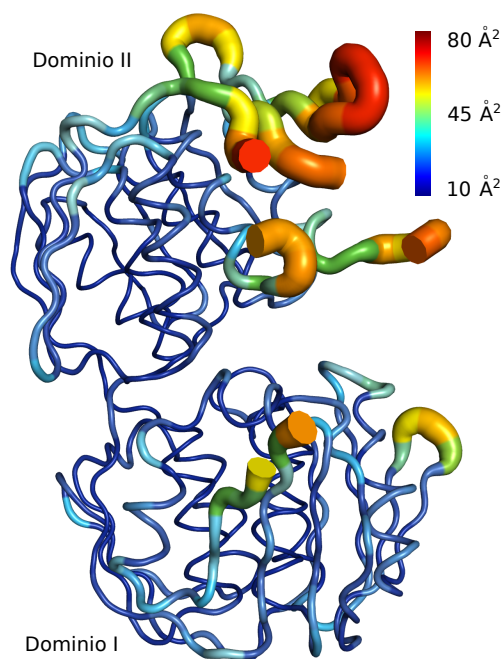
La proteína MurA1<sub>Hungary</sub><sup>19A</sup> cristalizó en el grupo ortorrómbico P2<sub>1</sub>2<sub>1</sub>2, con 4 moléculas por celdilla unidad. El eje  $c$  es binario simple mientras que  $a$  y  $b$  son binarios helicoidales. Al igual que sucedía en MurA1<sub>D39</sub>, cada dominio establece contactos con varios dominios de las moléculas vecinas. No obstante, los canales de solvente son más estrechos que en el cristal de MurA1<sub>D39</sub>, lo que implica una menor agitación y mayor resolución en la estructura final (Figura 4.8).



**Figura 4.8: Simetría y empaquetamiento de MurA1<sub>Hungary</sub><sup>19A</sup>.** El dominio I se representa en naranja, mientras que en azul se representa el dominio II. El recuadro negro hace referencia a la celdilla unidad, indicándose además la orientación de cada eje en cada caso. **a)** Plano  $(a, b)$ . **b)** Plano  $(b, c)$ . **c)** Plano  $(a, c)$ . **d)** Función de autorrotación calculada para  $\chi = 180^\circ$  mediante Molrep, mostrando los picos característicos de un grupo de Laue  $mmm$ .

### 4.8.3. Agitación térmica

Los factores B generales son sensiblemente más bajos que en el caso de MurA1<sub>D39</sub>, hecho directamente relacionado con la mayor resolución de los datos de MurA1<sub>Hungary</sub><sup>19A</sup>. Los factores B globales para la cadena principal se sitúan en una media de 18,1 Å<sup>2</sup>, mientras que el valor medio para las cadenas laterales es 23,9 Å<sup>2</sup>. Existen tres regiones con valores de agitación térmica significativamente superiores a la media (Pro82 – Ser90, Thr109 – Pro120 y Ala137 – Lys145), con valores que llegan a alcanzar los 80 Å<sup>2</sup>. Estas regiones se encuentran juntas en la misma zona del dominio II (Figura 4.9).

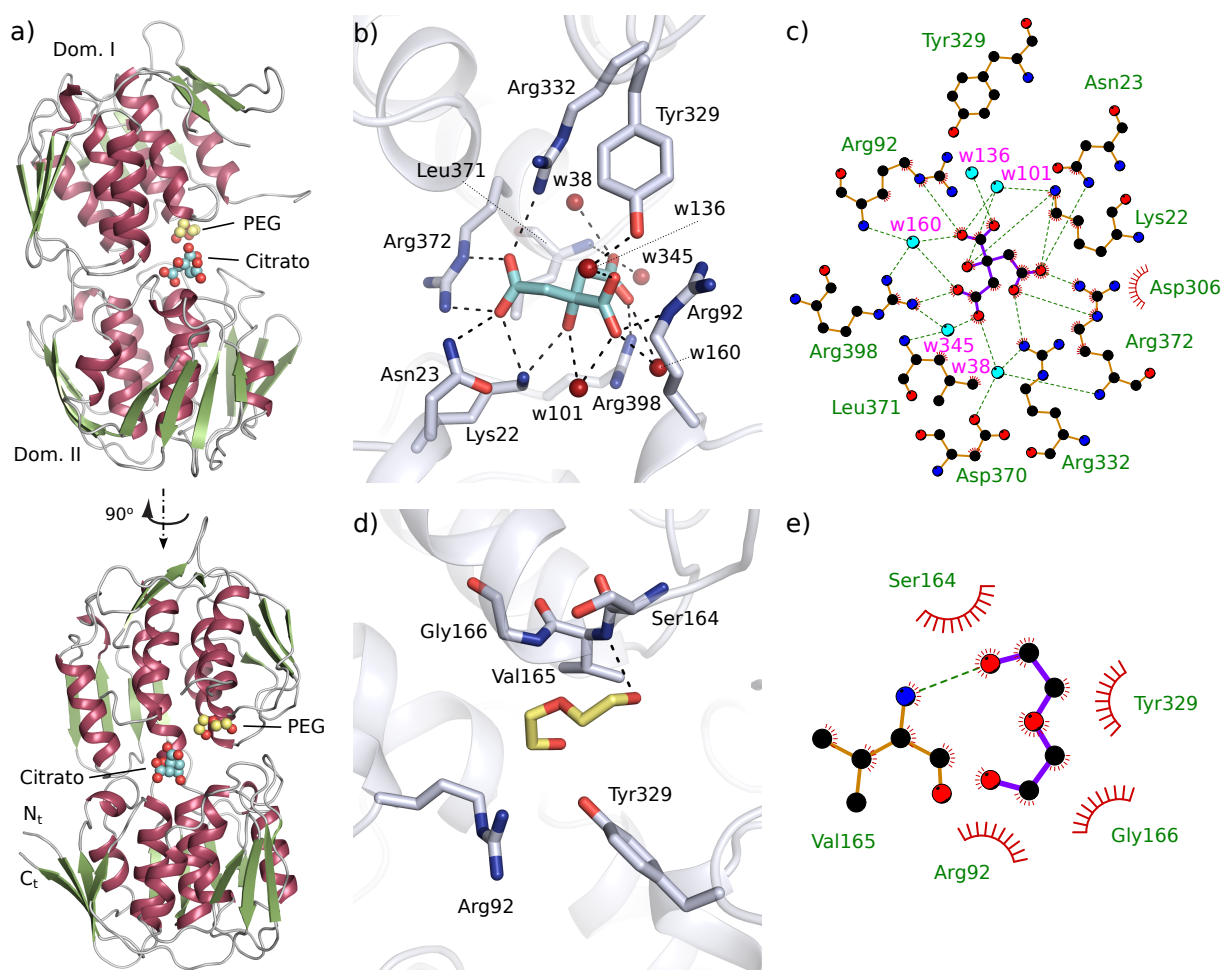


**Figura 4.9: Agitación térmica de MurA1<sub>Hungary19A</sub>.** Representación gráfica de la agitación térmica de MurA1<sub>Hungary19A</sub> en base a los factores B medios de cada residuo. Se puede observar una mayor agitación en las regiones carentes de modelización de la hélice  $\alpha_3$  (extremos del fragmento Pro82 – Ser90 entre 60 y 80  $\text{\AA}^2$ ) y del lazo flexible (extremos del fragmento Thr109 – Pro120 superiores a 60  $\text{\AA}^2$ ). Una tercera región, correspondiente con el lazo entre  $\beta_7$  y  $\beta_8$ , presenta mayor agitación, alcanzando los 80  $\text{\AA}^2$ . Otras zonas flexibles menos significativas se corresponden con lazos cortos expuestos al solvente (entre 40 y 60  $\text{\AA}^2$ ).

#### 4.8.4. Otras moléculas y solvente

La alta resolución de estos cristales permitió modelar hasta 348 moléculas de agua. Dos ligandos adicionales han sido identificados en el espacio entre los dos dominios de la proteína (Figura 4.10a). Uno de ellos es un citrato proveniente del tampón de pH utilizado durante la etapa de cristalización. Su factor B medio, situado en 27,2  $\text{\AA}^2$ , indica una fuerte estabilización fruto de las interacciones que se establecen con las cadenas laterales de la zona (Figuras 4.10b y 4.10c). La elevada cantidad de radicales hidroxilo que presenta el citrato le permite establecer interacciones directas con un total de 9 residuos (Lys22, Asn23, Arg92, Tyr329, Arg332, Leu371, Asp370, Arg372 y Arg398). Además, existen varias moléculas de agua que favorecen la estabilización del citrato tanto con las cadenas laterales cercanas como con otras más alejadas, mediante el establecimiento de una red de enlaces de hidrógeno. En concreto, 5 aguas son especialmente importantes por su cercanía al ligando (w38, w101, w136, w160 y w345).

La segunda molécula que se ha modelado es un PEG de 2 unidades, proveniente de la condición de cristalización, con un factor B medio de 45,6  $\text{\AA}^2$ , lo que indica una mayor agitación térmica con respecto a la estructura global. Este hecho se justifica con el establecimiento de menos interacciones que en el caso del citrato, ya que el PEG únicamente muestra una interacción fuerte de uno de sus hidroxilos terminales con el NH del enlace amida entre la Ser164 y la Val165. Existen no obstante otras interacciones más débiles que ayudan al mantenimiento de esta molécula en la posición que presenta, como la existente con el enlace amida contiguo al ya descrito o la estabilización parcial del otro extremo hidroxilo por parte de la Arg92 y la Tyr329 (Figuras 4.10d y 4.10e).



**Figura 4.10: Presencia de ligandos inespecíficos en MurA1<sub>Hungary19A</sub>.** a) Localización de las moléculas de citrato y PEG en el centro activo. b) Interacciones del citrato con los residuos más relevantes del sitio activo y con algunas moléculas de agua cercanas. c) Representación esquemática de las interacciones más importantes con el citrato (calculada con LigPlot<sup>+</sup>; Laskowski & Swindells, 2011). d) Interacciones de la molécula de PEG con los residuos más relevantes. Sólo existe una interacción fuerte entre este ligando y la enzima. e) Representación esquemática de las interacciones de la molécula de PEG con los residuos cercanos, mostrando la influencia de residuos alejados (calculada con LigPlot<sup>+</sup>).

#### 4.9. Cristalización de MurA1<sub>D39</sub> y MurA1<sub>Hungary19A</sub> en presencia de ligandos

Además de los ensayos de cristalización de las proteínas anteriormente descritos, se intentó obtener la estructura tanto de MurA1<sub>D39</sub> como de MurA1<sub>Hungary19A</sub> en complejo con distintos ligandos. Para ello se utilizaron las técnicas tanto de cocrystalización como de inmersión descritas en la Sección 3.2.3.

La inmersión se realizó sobre los cristales de MurA1<sub>D39</sub> y MurA1<sub>Hungary19A</sub> obtenidos en las etapas anteriormente descritas, mientras que las cocrystalizaciones se realizaron sobre un gradiente de condiciones de cristalización centradas en la condición de cristalización original óptima de cada proteína. De esta forma, si la solubilidad de la proteína se ve alterada por la presencia del ligando, se puede optimizar la cristalización en condiciones similares a la original. Además, debido a que la unión de los ligandos puede dar lugar a un cambio de conformación



(como una transición de la conformación abierta a la cerrada debida a la unión del ligando), se realizaron ensayos de cocrystalización de la MurA1<sub>D39</sub> en la condición de cristalización de la MurA1<sub>Hungary</sub><sup>19A</sup> y viceversa.

Uno de los ligandos utilizados en estos ensayos fue la UDP-*N*-acetilglucosamina (UDP-GlcNAc), por ser uno de los sustratos de la reacción. Al no añadir a la muestra fosfoenolpiruvato (PEP), el otro sustrato de la reacción catalizada por la MurA1, es de esperar que la UDP-*N*-acetilglucosamina se quede anclada en su sitio de unión sin ser procesada. Otro ligando que se ensayó fue la fosfomicina, antibiótico inhibidor de la MurA1 por unión al centro activo. Finalmente, se ensayaron los dos ligandos de manera conjunta con la intención de que primero se introdujera en el centro activo la UDP-*N*-acetilglucosamina y, posteriormente, la fosfomicina, mimetizando al fosfoenolpiruvato.

Se ensayaron diferentes concentraciones de ambos ligandos, tanto aislados como en conjunto, en un rango desde 2,5 mM hasta 50 mM, con tiempos de incubación en cocrystalización de 30 minutos y tiempos de inmersión desde pocos minutos hasta 24 horas. En ningún caso se consiguió cristalizar la proteína en complejo con alguno o ambos sustratos.

Por último, debido al cambio conformacional observado en MurA1<sub>Hungary</sub><sup>19A</sup> respecto a MurA1<sub>D39</sub>, se realizaron experimentos de cocrystalización e inmersión con citrato sobre la MurA1<sub>D39</sub>, sin llegar a conseguir cristales del complejo por ninguna de las técnicas. También se realizaron diversos ensayos de cristalización de MurA1<sub>Hungary</sub><sup>19A</sup> modificando el tampón de pH para conseguir la estructura libre de citrato, realizando al mismo tiempo gradientes de las concentraciones de los otros componentes y variando los puntos de pH, pero tampoco se llegaron a obtener cristales bajo estas condiciones.

## 4.10. Relación estructura-función

La función de la proteína MurA es bien conocida, consistente en catalizar la conversión de UDP-*N*-acetilglucosamina (UDP-GlcNAc) a enolpiruvil-UDP-*N*-acetilglucosamina (EP-UDP-GlcNAc) mediante la unión de fosfoenolpiruvato (PEP) y la liberación de fosfato inorgánico. Como se ha indicado previamente, las bacterias Gram positivas contienen dos genes, *murA1* y *murA2*, que codifican para dos proteínas MurA activas muy similares desde el punto de vista tanto de secuencia como enzimático (Du *et al.*, 2000). Además, estas proteínas están relacionadas con el fenómeno de heterorresistencia a fosfomicina en *S. pneumoniae*, como se verá a lo largo de esta sección.

### 4.10.1. Análisis comparado de estructuras

Como se ha indicado anteriormente, la única diferencia en secuencia entre MurA1<sub>D39</sub> y MurA1<sub>Hungary</sub><sup>19A</sup> es la sustitución Ala364Thr (Figura 4.11).

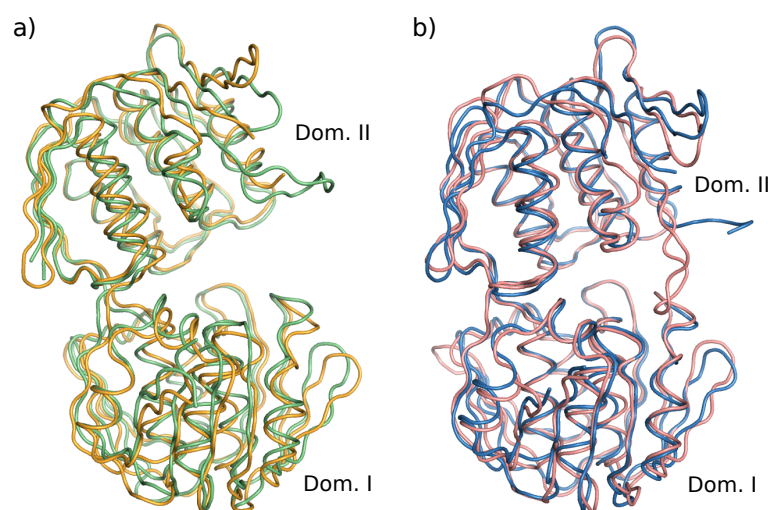
La proteína MurA1 de *S. pneumoniae* presenta un plegamiento similar al anteriormente descrito para las proteínas MurA de otras especies (Schönbrunn *et al.*, 1996; Skarzynski *et al.*, 1996). El homólogo estructural más cercano a MurA1<sub>D39</sub>, calculado mediante el servidor Dali

MurA1 <sub>D39</sub>	MRKIVINGGLPLQGEITITSGAKNSVVALIPAILLADDVVTLDCVPDISDV	50
MurA1 <sub>Hungary<sup>19A</sup></sub>	MRKIVINGGLPLQGEITITSGAKNSVVALIPAILLADDVVTLDCVPDISDV	50
*****		
MurA1 <sub>D39</sub>	ASLVEIMELMGATVKRYDDVLEIDPRGVQNI PMPYGKINS LRASYFYFGS	100
MurA1 <sub>Hungary<sup>19A</sup></sub>	ASLVEIMELMGATVKRYDDVLEIDPRGVQNI PMPYGKINS LRASYFYFGS	100
*****		
Lazo flexible		
MurA1 <sub>D39</sub>	LLGRFGEATVGLPGGCDLGPRPIDLHLKAFEAMGATASYEGDNMKLSAKD	150
MurA1 <sub>Hungary<sup>19A</sup></sub>	LLGRFGEATVGLPGGCDLGPRPIDLHLKAFEAMGATASYEGDNMKLSAKD	150
*****		
MurA1 <sub>D39</sub>	TGLHGASIYMDTVSVGATINTMIAAVKANGRTI IENAAAREPEI IDVATLL	200
MurA1 <sub>Hungary<sup>19A</sup></sub>	TGLHGASIYMDTVSVGATINTMIAAVKANGRTI IENAAAREPEI IDVATLL	200
*****		
MurA1 <sub>D39</sub>	NNMAHIRGAGTNI I I I IDGVERLHGTRHQV I PDRIEAGTYISLAAAVGKG	250
MurA1 <sub>Hungary<sup>19A</sup></sub>	NNMAHIRGAGTNI I I I IDGVERLHGTRHQV I PDRIEAGTYISLAAAVGKG	250
*****		
MurA1 <sub>D39</sub>	IRINNVLYEHLEGFIAKLEEMGVRMTVSEDSIFVEEQSNLKAINIKTAPY	300
MurA1 <sub>Hungary<sup>19A</sup></sub>	IRINNVLYEHLEGFIAKLEEMGVRMTVSEDSIFVEEQSNLKAINIKTAPY	300
*****		
MurA1 <sub>D39</sub>	PGFATDLQQPLTPLLRRANGRTIVDTIYEKRVNHVFELAKMDADISTTN	350
MurA1 <sub>Hungary<sup>19A</sup></sub>	PGFATDLQQPLTPLLRRANGRTIVDTIYEKRVNHVFELAKMDADISTTN	350
*****		
MurA1 <sub>D39</sub>	GHILYTGGRDLRGASVKATDLRAGAALVIAGLMAEGKTEITNIEFILRGY	400
MurA1 <sub>Hungary<sup>19A</sup></sub>	GHILYTGGRDLRGTSVKATDLRAGAALVIAGLMAEGKTEITNIEFILRGY	400
*****		
MurA1 <sub>D39</sub>	SDIEKLRLNLGADIRLVED	419
MurA1 <sub>Hungary<sup>19A</sup></sub>	SDIEKLRLNLGADIRLVED	419
*****		

**Figura 4.11:** Alineamiento de secuencias de MurA1<sub>D39</sub> y MurA1<sub>Hungary<sup>19A</sup></sub>. En naranja se representa el dominio I, mientras que en azul se ha representado en azul. En negro han quedado representados los residuos involucrados en la bisagra doble. El recuadro rojo hace referencia a la Cys116 catalítica, mientras que el recuadro azul señala la sustitución Ala364Thr.

(Holm & Rosenström, 2010) se corresponde a la conformación abierta de la proteína MurA de *E. cloacae* (código PDB 1EJD, RMSD = 1,8 Å; Eschenburg & Schönbrunn, 2000) (Figura 4.12a). En cambio, el homólogo estructural más cercano a MurA1<sub>Hungary<sup>19A</sup></sub> es la proteína MurA de *Bacillus anthracis* (código PDB 3SG1, RMSD = 1,3 Å; sin publicación asociada), que presenta una conformación cerrada (Figura 4.12). Así pues, nos encontramos frente a dos conformaciones, abierta y cerrada, de la proteína MurA1 de *S. pneumoniae*, como se podía deducir a partir del comportamiento del reemplazo molecular y de la diferencia en la separación angular entre los dominios I y II de la proteína en cada caso (Sección 4.8.1).

El lazo flexible que contiene la cisteína catalítica (Cys116) consta de 10 residuos bastante conservados entre las proteínas MurA de diferentes especies (Figura 4.13). Las prolinas de los extremos aportan la flexibilidad al lazo, actuando como bisagras que acercan y alejan la Cys116 del sitio catalítico. Las posiciones 117 y 120, ocupadas en *S. pneumoniae* por un ácido aspártico y una prolina respectivamente, no aparecen conservadas en ninguna otra estructura de MurA, las cuales generalmente contienen alaninas, treoninas o serinas. Por último, es especialmente interesante el caso de la proteína MurA de *Mycobacterium tuberculosis*, que presenta un ácido



**Figura 4.12: Alineamiento estructural de MurA1.** a) Alineamiento entre MurA1<sub>D39</sub> (naranja) y la conformación abierta de la proteína MurA de *E. cloacae* (verde). b) Alineamiento entre MurA1<sub>Hungary19A</sub> (azul) y la conformación cerrada de MurA de *B. anthracis* (rosa).

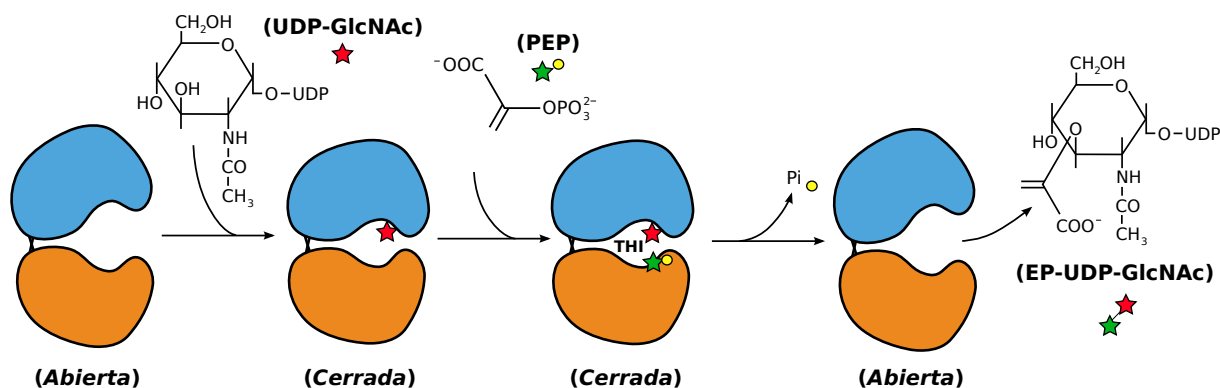
aspártico en lugar de la Cys116. Como se ha observado en otros trabajos, la sustitución de la cisteína catalítica por un aspártico relaja la regulación de la formación de producto y hace a la enzima más vulnerable frente a mayor número de antibióticos de bajo peso molecular (Zhu *et al.*, 2012).

#### 4.10.2. Reconocimiento de sustratos y sitio activo

Ha sido previamente publicado que el mecanismo de acción de la enzima MurA se inicia con la proteína en conformación abierta, recibiendo secuencialmente sus dos sustratos, a saber, UDP-GlcNAc y PEP. La entrada de la UDP-GlcNAc produce un cambio de conformación en la proteína, cerrándose sobre ella, acercando sus dominios y ubicando la cisteína catalítica en una posición adecuada para que se inicie la reacción de transferencia del grupo enolpiruvil del PEP (unido tras el cierre conformacional) a la UDP-GlcNAc, dando lugar al EP-UDP-GlcNAc (Cassidy & Kahan, 1973; Han *et al.*, 2010). El producto de la reacción se libera mediante la apertura de la estructura, adquiriendo nuevamente la conformación abierta. Además, durante el transcurso de la reacción, se forma un intermediario tetrahédrico (THI) cuya estabilidad es crucial para el transcurso de la reacción hacia la formación del producto (Figura 4.14).

	113	*	122
<i>Streptococcus pneumoniae</i>	PGGC	DLG	PRP
<i>Enterobacter cloacae</i>	PGGCA	IGARP	
<i>Bacillus anthracis</i>	PGGCA	IGSRP	
<i>Escherichia coli</i>	PGGCT	IGARP	
<i>Vibrio fischeri</i>	PGGCA	IGARP	
<i>Haemophilus influenzae</i>	PGGCS	IGARP	
<i>Listeria monocytogenes</i>	PGGCA	IGSRP	
<i>Aquifex aeolicus</i>	PGGCS	IGARP	
<i>Mycobacterium tuberculosis</i>	PGGDA	IGSRP	

**Figura 4.13: Alineamiento de los lazos flexibles de MurA en diferentes especies.** La Cys116 (\*) aparece conservada en todas las especies salvo en *M. tuberculosis*. Las posiciones 117 y 120 presentan una mayor variabilidad (recuadros).



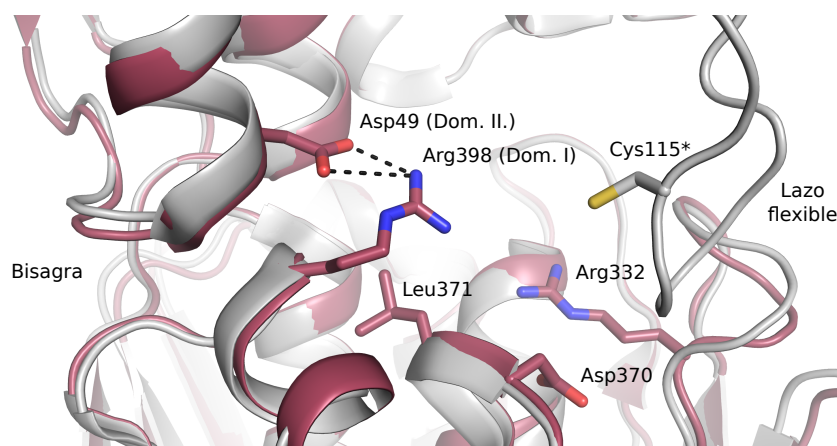
**Figura 4.14: Esquema del mecanismo de reacción de MurA1.** La conformación abierta de MurA1 recibe el primer sustrato, la UDP-GlcNAc. Su entrada en el sitio activo provoca el cierre de la estructura. Posteriormente, el PEP entra en su sitio y, tras formar un intermediario tetrahédrico (THI), se genera EP-UDP-GlcNAc, liberándose fosfato inorgánico ( $P_i$ ). Finalmente, la estructura de MurA1 recupera la conformación abierta inicial.

Si bien no ha sido posible obtener una estructura de MurA1 de *S. pneumoniae* que contenga alguno de los sustratos de la reacción, se puede realizar una extrapolación de los residuos implicados no sólo en la catálisis sino también en el reconocimiento y estabilización de los sustratos a partir de la superposición y análisis comparado de MurA1<sub>D39</sub> y MurA1<sub>Hungary19A</sub> con la proteína MurA de otros microorganismos en conformación abierta y cerrada respectivamente. En concreto, además de las proteínas homólogas estructurales utilizadas con anterioridad, se ha estudiado principalmente el sitio catalítico de la MurA de *E. coli* resuelta por Skarzynski *et al.*, 1996, que contiene UDP-GlcNAc y fosfomicina (código PDB 1UAE). Además, existen otras estructuras cuyo análisis comparado resulta interesante debido a la presencia de mutaciones que bloquean alguna etapa de la reacción. En *E. coli* se ha resuelto la mutación Cys<sub>115</sub>Ala (código PDB 1A2N; Skarzynski *et al.*, 1998), que bloquea el THI y mantiene la estructura cerrada y en interacción con los sustratos reales de la enzima. El mismo efecto se consiguió en la MurA de *E. cloacae* mediante la sustitución Asp<sub>305</sub>Ala (código PDB 1Q3G; Eschenburg *et al.*, 2003). Por último, otra mutación interesante es la realizada por Eschenburg *et al.*, 2005, también en la MurA de *E. cloacae* (código PDB 1RYW), que bloquea el producto de la reacción en el sitio catalítico.

Como propuso Jackson *et al.*, 2009, y como se deduce de los análisis comparados con la proteína MurA1 de *S. pneumoniae*, los residuos Asp49 y Arg398 podrían estar implicados en el proceso de transición de la conformación abierta a la cerrada y viceversa, al mismo tiempo que podrían desarrollar algún papel en la liberación del fosfato inorgánico, subproducto de la reacción. Este hecho podría ser el desencadenante de una serie de movimientos a lo largo de los residuos Leu371, Asp370, Arg332 y Glu330 que terminasen en la liberación del lazo flexible, abriendo nuevamente la estructura (Figura 4.15).

En lo que respecta a la unión de la UDP-GlcNAc, el anillo de uridina podría localizarse entre dos superficies hidrofóbicas generadas por los residuos Arg121 – Pro122 y por la Leu125 respectivamente, como sugiere Skarzynski *et al.*, 1996. Además, otros aminoácidos que podrían estar implicados en el reconocimiento del UDP son la Arg92, Asp124, Ser164, Val165, Gly166 y





**Figura 4.15: Interacciones implicadas en el cambio conformacional de MurA1.**

Las interacciones existentes entre el Asp49 del dominio II y la Arg398 del dominio I en MurA1<sub>Hungary</sub><sup>19A</sup> (representada en rojo) mantienen la conformación cerrada. En MurA1<sub>D39</sub>, de conformación abierta, estos dos residuos se encuentran separados por una distancia de 10 Å y no interactúan entre sí. La posición de la Arg398 podría desplegar una serie de cambios en la orientación de los residuos vecinos que se prolongase hasta el lazo flexible, bloqueándolo en la posición cerrada (estructura blanca, correspondiente a la proteína MurA de *E. cloacae* D305A, código PDB 1Q3G; Eschenburg *et al.*, 2003). La Cys115 se ha etiquetado con \* por seguir una numeración diferente a la de MurA1<sub>Hungary</sub><sup>19A</sup>.

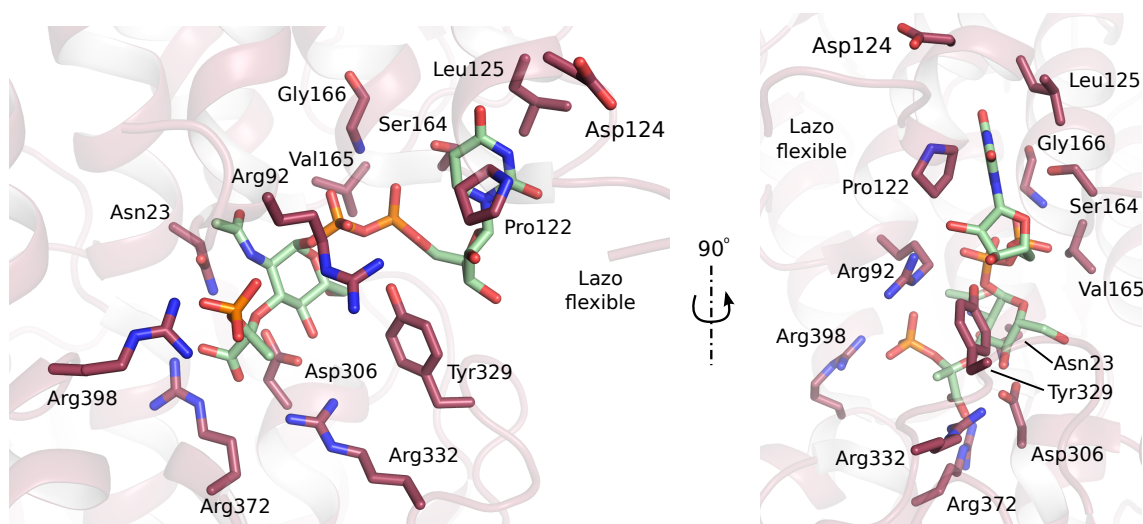
Tyr329 (Figura 4.16). Por otro lado, el fragmento de GlcNAc podría interactuar con la Asn23 y el Asp306, formando puentes de hidrógeno con el hidroxilo 3' del sustrato.

Por último cabe destacar que, debido a la ausencia de estructuras que contengan PEP en el centro catalítico, no es tarea sencilla determinar qué residuos de MurA1 pueden estar implicados en su reconocimiento y unión. No obstante, como se indicó anteriormente, existen mutantes que bloquean el THI en el sitio catalítico y permiten analizar qué residuos están en contacto con la zona del intermediario correspondiente al PEP (Eschenburg *et al.*, 2003). De este modo se puede apuntar a la Arg372 como candidato más probable a estabilizar el PEP durante la reacción. Además, basándose una vez más en el trabajo de cristalización de MurA en presencia de (Z)-3-fluorofosfoenolpiruvato (análogo no hidrolizable de PEP) de Skarzynski *et al.*, 1998, se puede inferir que los residuos Arg332 y Arg398 podrían participar en la unión de PEP (Figura 4.16).

Pese a conocer diversas estructuras de la proteína MurA, en conformaciones abierta y cerrada y en presencia y ausencia de diferentes sustratos, el mecanismo de reacción no se conoce en detalle. Si bien se ha podido confirmar que la unión de la UDP-GlcNAc sucede en primer lugar (Cassidy & Kahan, 1973), aún no queda claro si el PEP se une covalentemente a algún aminoácido del sitio activo o si es una especie libre la que reacciona con la UDP-GlcNAc.

#### 4.10.3. Inhibición mediante fosfomicina y citrato

Como se introdujo en la Sección 1.5.1, la proteína MurA es inhibida por fosfomicina. La fosfomicina es un antibiótico de amplio espectro, derivado del ácido fosfónico, que presenta una elevada avidez por grupos sulfhidrilos expuestos, estableciendo enlaces covalentes irreversibles con ellos (Hendlin *et al.*, 1969; Raz, 2012). Por tanto, es muy probable que este antibiótico



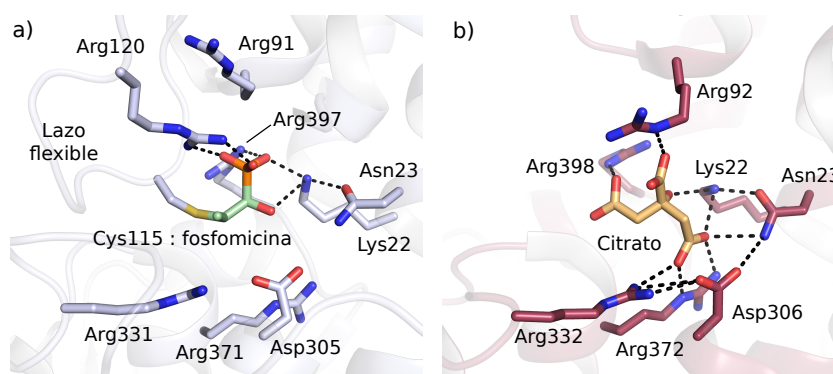
**Figura 4.16: Residuos implicados en el reconocimiento de sustratos de MurA1.** El intermediario tetrathédrico (en verde) pertenece a la proteína MurA de *E. cloacae* D305A (código PDB 1Q3G; Eschenburg *et al.*, 2003). La superposición de esta estructura con MurA1<sub>Hungary19A</sub> (en rojo) permite analizar el sitio de unión tanto a PEP como a UDP-GlcNAc. Se han representado los residuos que, bien por la información previa, bien por la cercanía con el sustrato, pueden estar implicados en el reconocimiento de los sustratos.

aparezca unido a la estructura si se ha añadido durante los experimentos de cristalización. No obstante, si no se logra la estabilización del lazo flexible que contiene la Cys116, no se observará su unión. La estructura resuelta por Skarzynski *et al.*, 1996, muestra la existencia de un enlace de hidrógeno entre la fosfomicina unida a la cisteína catalítica y el 3'-OH de la UDP-GlcNAc. A partir del trabajo de Han *et al.*, 2010, se pueden determinar los residuos bloqueados por la acción directa de la fosfomicina (Figura 4.17a).

En el caso de MurA1 de *S. pneumoniae*, no se pudo observar el complejo MurA1 : fosfomicina. Partiendo de la hipótesis de que la fosfomicina se ha unido a la Cys116 durante la etapa de incubación, la elevada flexibilidad del lazo que contiene a la Cys116 impide observar el complejo. Además, en el caso de MurA1<sub>Hungary19A</sub>, la presencia de la molécula de citrato en el sitio de unión de PEP/fosfomicina puede impedir la estabilización de la fosfomicina entre los dos dominios estructurales, evitando nuevamente su observación (Figura 4.17b).

La comparación entre el sitio de unión a citrato observado en MurA1<sub>Hungary19A</sub> y el sitio de unión a PEP/fosfomicina de Han *et al.*, 2010, refuerza la hipótesis de que el citrato funciona como análogo estructural tanto del PEP como de la fosfomicina, insertándose en su posición. El citrato se estabiliza con mayor número de interacciones que el PEP o la fosfomicina. Así pues, el citrato podría desplazar al PEP y bloquear la enzima de forma aparentemente irreversible. Además, el hecho de que establezca una unión más fuerte que la fosfomicina y que se haya añadido a la condición de cristalización antes que la fosfomicina justifica el motivo por el que no se observa ni la fosfomicina ni el lazo flexible estabilizado en la estructura final de MurA1<sub>Hungary19A</sub>.

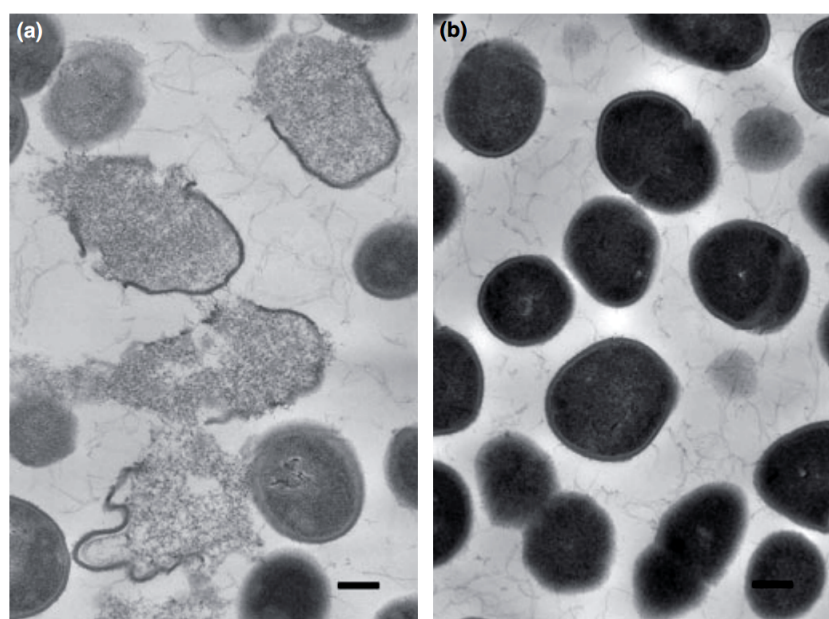
Por último, como se ha comentado previamente, el mecanismo propuesto para MurA implica el cierre de la estructura mediante la entrada del UDP-GlcNAc como primer sustrato (Figura 4.14). En cambio, la estructura de MurA1<sub>Hungary19A</sub> presenta una conformación cerrada en



**Figura 4.17: Comparación entre los sitios de unión de fosfomicina y citrato en MurA.** a) Detalle de la inhibición por fosfomicina en la proteína MurA de *E. cloacae* (código PDB 3LTH; Han *et al.*, 2010). Se observa la unión covalente entre la Cys115 y la fosfomicina, que bloquea a los residuos Lys22, Asn23 y Arg397. b) Sitio de reconocimiento de citrato en MurA1<sub>Hungary19A</sub>. Además de los residuos implicados en la unión a fosfomicina, se establecen interacciones con los residuos Arg92, Asp306, Arg332 y Arg372.

ausencia de UDP-GlcNAc pero en presencia de citrato. Esto se debe a que el citrato establece enlaces tanto con residuos del sitio de unión a PEP/fosfomicina (ubicados en el dominio II) como con residuos del dominio I que estabilizan parte de la UDP-GlcNAc (Figura 4.15b). En otras palabras, el citrato actúa como puente entre los dos dominios, acercándolos entre sí y cerrando la estructura. De este modo no sólo se bloquea el sitio de unión a PEP/fosfomicina, sino que también se impide la entrada de UDP-GlcNAc en el sitio catalítico.

El hecho de que el citrato haya inhibido fuertemente a MurA1 sugiere su posible utilización, o la de un análogo estructural, como agente antimicrobiano a nivel terapéutico. Ensayos de viabilidad celular realizados sobre *S. pneumoniae* en presencia de citrato sódico han mostrado la eficacia de este compuesto como agente bacteriolítico (Figura 4.18) (Nagaoka *et al.*, 2010).

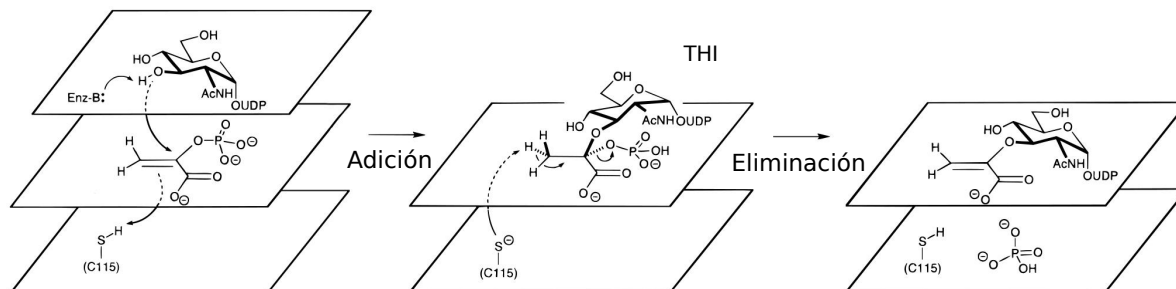


**Figura 4.18: Efecto del citrato sódico en células de *S. pneumoniae*.** Fotografías de TEM (Microscopía Electrónica de Transmisión), a x20000, de células de *S. pneumoniae*. a) Cultivo en presencia de 25 mmol/L de citrato sódico. b) Cultivo en ausencia de citrato sódico. La escala hace referencia a 200 nm. Figura extraída de Nagaoka *et al.*, 2010.

## 4.10.4. Hipótesis sobre el mecanismo de reacción

Existen actualmente tres hipótesis que explican el mecanismo de reacción en la proteína MurA de otros microorganismos y que también son aplicables a la MurA1 de *S. pneumoniae*, ya que los residuos implicados en cada una aparecen conservados entre todas las estructuras resueltas:

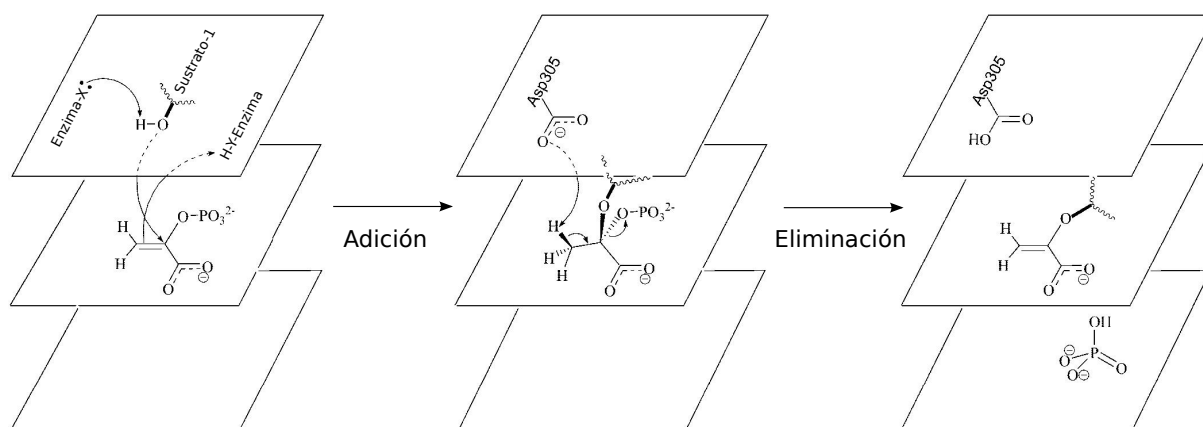
1. La Cys116 actúa como catalizador ácido general, llevando a cabo el paso de adición de la etapa de adición-eliminación, mientras que la UDP-GlcNAc lleva a cabo el ataque nucleofílico. En este caso, el PEP puede unirse covalentemente a la Cys116 (Brown *et al.*, 1994) o puede ser introducido en el sitio catalítico sin necesidad de que exista unión covalente alguna (Kim *et al.*, 1996; Skarzynski *et al.*, 1998). En este último caso, la Cys116 no colocaría el PEP cerca del 3'-OH de la UDP-GlcNAc, sino que iniciaría la transferencia del protón que daría lugar a la formación del THI. A continuación, debido al paso de eliminación de la etapa de adición-eliminación, el THI se disgregaría también por la acción de la Cys116, desprotonando el C<sub>3</sub> del fragmento correspondiente al PEP y permitiendo la liberación del producto principal, EP-UDP-GlcNAc, y fosfato inorgánico (Figura 4.19). Bajo esta hipótesis, la inhibición de la MurA1 por fosfomicina podría explicarse mediante una unión covalente entre el antibiótico y la Cys116, bloqueando bien la unión del PEP, bien la transferencia del protón en la fase de adición (en función de si se considera o no la unión covalente entre el PEP y la Cys116).



**Figura 4.19: Mecanismo de reacción de MurA1 basado en la cisteína catalítica.**

Figura adaptada de Skarzynski *et al.*, 1998.

2. La Lys22 es el residuo que provoca las dos transferencias de protones requeridas en el paso de adición de la reacción, mientras que el Asp306 estaría implicado en la desprotonación del fragmento correspondiente al PEP en el THI durante la fase de eliminación (Eschenburg *et al.*, 2003) (Figura 4.20). No obstante, Mizyed *et al.*, 2003, ha sugerido que no existe una cadena lateral que catalice sólo la adición o la eliminación, lo que contradice esta hipótesis. Como consecuencia, la Cys116 no participaría en la reacción. Ya que las estructuras inhibidas por fosfomicina sólo contienen UDP-GlcNAc, se puede inferir que la Cys116 juega un papel importante en el reconocimiento del PEP y en su orientación dentro del sitio activo, pero no participa directamente en la reacción.
3. Se produce una ruptura no enzimática del THI provocada por la liberación del fosfato inorgánico tras el corte de un enlace C-O (Byczynski *et al.*, 2003) (Figura 4.21). Aunque

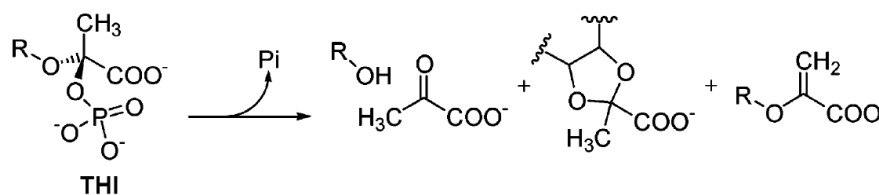


**Figura 4.20: Mecanismo de reacción de MurA1 sin mediación de la cisteína.** Figura adaptada de Eschenburg *et al.*, 2003.

este corte es producido por un catalizador ácido general intramolecular, la Lys22 podría apoyar indirectamente este proceso actuando como catalizador ácido general externo. En lo que respecta a la inhibición, la fosfomicina jugaría el mismo papel que en la hipótesis anterior.

#### 4.10.5. MurA1<sup>Hungary<sup>19A</sup></sup> y heterorresistencia a fosfomicina

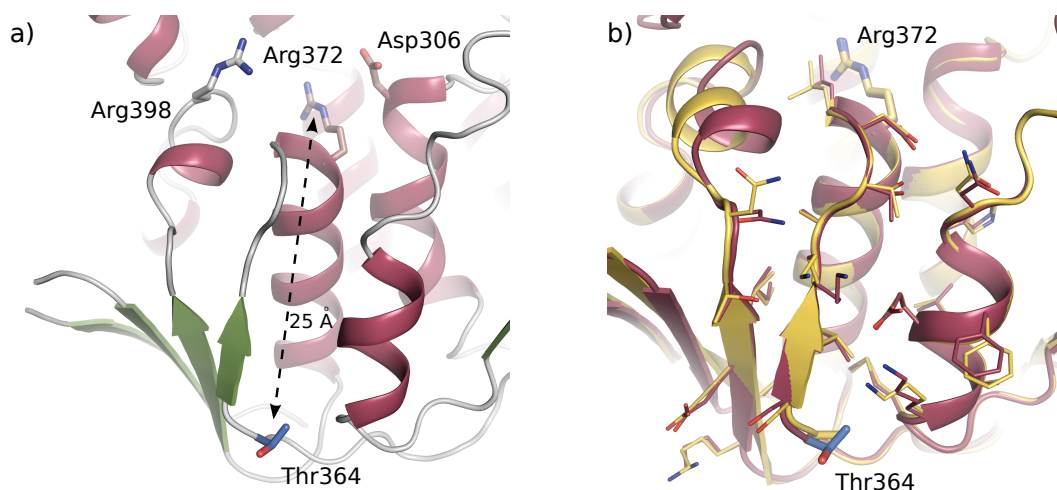
Trabajos realizados en el grupo de la Prof. Kathrin Mühlemann y la Dra. Lucy J. Hathaway (Instituto de Enfermedades Infecciosas, Facultad de Medicina, Universidad de Berna, Suiza) han demostrado la inhibición tanto de MurA1 como de MurA2 por fosfomicina. Además, de 11 cepas probadas, todas mostraron heterorresistencia a fosfomicina salvo una, a saber, la cepa Hungary<sup>19A</sup>. En cambio, sólo 8 cepas mostraron heterorresistencia a penicilina, por lo que se espera que los mecanismos de heterorresistencia de ambos antibióticos difieran entre sí. Mientras que los ensayos microbiológicos de crecimiento en presencia de fosfomicina mostraban una distribución de concentraciones mínimas inhibitorias (CMI) habitual (crecimiento bajo concentraciones de fosfomicina de 3 a 32  $\mu\text{g/mL}$ ), las cepas heterorresistentes generaban subpoblaciones capaces de crecer a una concentración de fosfomicina de 300  $\mu\text{g/mL}$ . En base a este resultado, se analizaron los elementos diferenciales entre la cepa Hungary<sup>19A</sup> y las cepas heterorresistentes, encontrándose únicamente la mutación Ala364Thr en la proteína MurA1. Así pues, se realizaron experimentos adicionales para analizar el papel de MurA1 en el fenómeno de heterorresistencia a fosfomicina.



**Figura 4.21: Mecanismo de reacción de MurA1 con ruptura espontánea del THI.** La ruptura espontánea del THI lleva a la formación de diferentes productos. Este mecanismo podría estar favorecido por la acción de algún residuo cercano. Figura adaptada de Byczynski *et al.*, 2003.

En primer lugar, se observó que la delección de *murA1* eliminaba la heterorresistencia a fosfomicina, manteniendo invariante la heterorresistencia a penicilina. Esta mutación no es deletérea ya que, en ausencia de MurA1, es MurA2 la que inicia la síntesis de peptidoglicano, generando células viables. Además, se analizaron los niveles de expresión tanto de *murA1* como de *murA2* en diferentes cepas, sin que se observaran discrepancias. Así pues, se descarta que la heterorresistencia a fosfomicina se deba a una expresión diferencial de alguno de estos genes.

Por tanto, la existencia de la mutación puntual Ala364Thr entre la MurA1 de las cepas heterorresistentes y la cepa Hungary<sup>19A</sup> podría ser la causa del fenómeno. Como se ha visto en las secciones anteriores, la Ala364 no es uno de los residuos implicados en la formación del sitio activo ni de los residuos cercanos que pueden crear un ambiente idóneo para la ejecución de la reacción. Además, su lejanía respecto al centro activo ( $\simeq 25$  Å) imposibilita su implicación directa en el mecanismo de reacción (Figura 4.22a).



**Figura 4.22: Localización de Ala364Thr y organización del entorno. a)** Distancia entre el centro activo y la ubicación de la Thr364 en MurA1<sub>Hungary19A</sub>. **b)** Superposición de MurA1<sub>Hungary19A</sub> (morado) y MurA1<sub>D39</sub> (amarillo), mostrando los residuos cercanos a la mutación puntual y al centro activo. No se observan cambios significativos en la conformación de ninguna cadena que impliquen un cambio en la estructura del sitio activo de MurA1<sub>Hungary19A</sub> respecto a MurA1<sub>D39</sub>.

Por otro lado, la mutación se encuentra al inicio de la  $\beta_{21}$ , que lleva directamente a la Arg372, posiblemente implicada en el reconocimiento del PEP. Una sustitución de Ala por Thr añadiría un grupo hidroxilo a una región en la que existen numerosos puentes salinos entre las cadenas laterales cercanas, por lo que se podría producir un cambio en la estructura de esta red de interacción, alterando el plegamiento de la zona y desplazando, en último término, la Arg372 de su posición en el sitio activo. Como consecuencia, las argininas 306 y 332 también podrían sufrir una reorientación indirecta, afectando por tanto a una zona considerable del sitio activo (Figura 4.22a). No obstante, esta hipótesis ha quedado descartada al analizar estas regiones en MurA1<sub>Hungary19A</sub>, ya que se mantienen invariantes respecto a MurA1<sub>D39</sub> pese a la inclusión del hidroxilo de la Thr364 (RMSD = 0,546 para el dominio I; Figura 4.22b).

Por tanto, no existen evidencias sólidas que indiquen que la falta de heterorresistencia a fosfomicina en *S. pneumoniae* Hungary<sup>19A</sup> se deba a causa estructural. Para reforzar esta



hipótesis, se realizó una transformación del gen *murA1*<sub>Hungary<sup>19A</sup></sub> en la cepa D39 $\Delta$ *murA1*, con la intención de analizar el comportamiento de una cepa heterorresistente que expresara el gen *murA1* de una cepa no heterorresistente. Este experimento no cambió el fenotipo de la cepa D39, que se mantuvo heterorresistente. También se realizó el experimento inverso, introduciendo el gen *murA1*<sub>D39</sub> en la cepa Hungary<sup>19A</sup> $\Delta$ *murA1*, sin que apareciese heterorresistencia a fosfomicina. Así pues, el intercambio de los genes de *murA1* no lleva a un intercambio de fenotipos entre cepas. Finalmente, dado que el producto de reacción de MurB (UDP-MurNAC) puede ejercer una regulación negativa sobre MurA (Mizyed *et al.*, 2005; Zhu *et al.*, 2012), se realizaron experimentos similares a los anteriormente descritos cruzando el gen *murB* junto con el correspondiente gen *murA1*. En ningún caso el fenotipo de heterorresistencia a fosfomicina se vio alterado.

Así pues, con esta serie de experimentos tanto microbiológicos como estructurales no sólo se demuestra que la proteína MurA1 es necesaria para la existencia de heterorresistencia a fosfomicina, sino también que la causa de este fenómeno no está relacionada con su estructura. Por tanto, debe existir otro factor que influya en el mecanismo de inhibición de fosfomicina, como se ha descrito en el caso de heterorresistencia a metilicina por *S. aureus* (Murakami & Tomasz, 1989).





---

# DETERMINACIÓN ESTRUCTURAL Y FUNCIONAL DE DacB

---

## Índice

5.1	Caracterización de DacB . . . . .	<b>100</b>
5.2	Cristalización y resolución parcial de DacB <sub>20–238</sub> . . . . .	<b>100</b>
5.2.1	Expresión y purificación . . . . .	100
5.2.2	Cristalización y optimización . . . . .	101
5.2.3	Determinación estructural . . . . .	102
5.2.4	Análisis estructural . . . . .	105
5.3	Segunda construcción (DacB <sub>61–238</sub> ) . . . . .	<b>109</b>
5.3.1	Expresión y purificación . . . . .	109
5.3.2	Cristalización y resultados preliminares . . . . .	109
5.3.3	Variantes de la cristalización para evitar maclas . . . . .	113
5.3.4	Análisis de maclas y determinación estructural . . . . .	116
5.3.5	Faseado experimental de cristales maclados . . . . .	122
5.3.6	Determinación de la existencia del sitio de unión a zinc . . . . .	124
5.3.7	Análisis estructural . . . . .	124
5.4	Tercera construcción (DacB <sub>49–238</sub> ) . . . . .	<b>127</b>
5.4.1	Expresión y purificación . . . . .	127
5.4.2	Cristalización . . . . .	128
5.4.3	Determinación estructural . . . . .	128
5.4.4	Análisis estructural . . . . .	130
5.4.5	Relación entre DacB <sub>49–238</sub> y DacB <sub>61–238</sub> . . . . .	135
5.4.6	Cristalización en presencia de ligandos . . . . .	137
5.5	Relación estructura-función . . . . .	<b>138</b>
5.5.1	Maduración de DacB . . . . .	138
5.5.2	Análisis estructural comparado . . . . .	139
5.5.3	Maquinaria catalítica y conformaciones estructurales . . . . .	141
5.5.4	Estudios microbiológicos y fisiológicos relacionados con DacA y DacB .	145

---

## 5.1. Caracterización de DacB

El objetivo de los experimentos descritos en las siguientes secciones es la determinación estructural y funcional de la proteína DacB de *S. pneumoniae* TIGR4.

La proteína DacB de *S. pneumoniae* consta de 238 aminoácidos, con un peso teórico de 26,5 kDa y un pI teórico de 5,3. El análisis de su secuencia señala que DacB se divide en una región N-terminal (residuos Met1 a Ala55) y en un dominio catalítico (residuos Ala55 a Asp238) (Figura 5.1). La región N-terminal constaría del péptido señal característico de las lipoproteínas (residuos Tyr1 a Cys19) que contiene el motivo *LipoBox*, seguido de una zona flexible que llegaría hasta el residuo Ala55. Dicha zona presenta una probabilidad de desorden superior al 70 % según DISOPRED2 (Linding *et al.*, 2003), PONDRFIT (Ward *et al.*, 2004) y DisEMBL (Xue *et al.*, 2010).

```

      10          20          30          40          50          60
MKKRYLVLTAL LLAALS LAACS QEKTKNEDGE TKTEQTAKAD GTVGSKSQGA AQKKAEVVNK
      70          80          90         100         110         120
GDYYSIQGKY DEIIVANKHY PLSKDYNPGE NPTAKAELVK LIKAMQEAGF PISDHYSGFR
     130         140         150         160         170         180
SYETQTKLYQ DYVNQDGKAA ADRYSARPGY SEHQTGLAFD VIGTDGDLVT EEKAAQWLLD
     190         200         210         220         230
HAADYGFVVR YLKGKEKETG YMAEEWHLRY VGKEAKEIAA SGLSLEEYYG FEGGDYVD

```

**Figura 5.1:** Secuencia de la proteína DacB de *S. pneumoniae* TIGR4. En gris se muestra el péptido señal, mientras que en rojo se muestra el dominio catalítico. El recuadro azul se sitúa sobre el *LipoBox*.

Además, existía la sospecha de la existencia de un sitio de unión a zinc. Si bien en el momento de realizar el análisis de secuencia no existían evidencias concluyentes que confirmasen esta hipótesis, los predictores sugerían la existencia de un patrón habitual de metaloenzimas que unen zinc.

## 5.2. Cristalización y resolución parcial de DacB<sub>20–238</sub>

Para resolver la estructura tridimensional de DacB se procedió, en primer lugar, a la cristalización de la proteína completa sin el péptido señal (residuos 1 al 19). Así pues, el fragmento ensayado cubría los residuos Ser20 – Asp238 (DacB<sub>20–238</sub>; Figura 5.2).

### 5.2.1. Expresión y purificación

La obtención de DacB<sub>20–238</sub> fue realizada por el grupo del Prof. Sven Hammerschmidt, del Departamento de Genética Microbiana de la Universidad de Greifswald, Alemania.

Las secuencias de *dacB* se amplificaron mediante PCR usando DNA cromosómico de *S. pneumoniae* TIGR4 como molde. Los cebadores incluyeron los sitios de restricción *NheI* y

10	20	30	40	50	60
MKKRYLVLT	LLALS	LAACS	QEKTKNEDGE	TKTEQTAKAD	GTVGSKSQGA
AQKKA	AEVVNK				
70	80	90	100	110	120
GDYYSIQGKY	DEIIVANKHY	PLSKDYNPGE	NPTAKAELVK	LIKAMQEAGF	PISDHYSGFR
130	140	150	160	170	180
SYETQTKLYQ	DYVNQDGKAA	ADRYSARPGY	SEHQTGLAFD	VIGTDGDLVT	EEKAAQWLLD
190	200	210	220	230	
HAADYGFVVR	YLGKKEKETG	YMAEEWHLRY	VGKEAKEIAA	SGLSLEEYYG	FEGGDYVD

**Figura 5.2: Secuencia de la construcción DacB<sub>20–238</sub>** En gris se muestra el péptido señal, mientras que en rojo se muestra el dominio catalítico. El sombreado amarillo se corresponde con la secuencia de la construcción de DacB<sub>20–238</sub>.

*Hind*III para la posterior inclusión de la construcción en el plásmido pTP1 (Saleh *et al.*, 2013). El plásmido resultante se introdujo en *E. coli* BL21(DE3).

Para la producción de DacB<sub>20–238</sub> se cultivó la cepa de *E. coli* transformada con el plásmido en medio líquido LB suplementado con kanamicina (50 µg/mL) a una temperatura de 30 °C hasta alcanzar una A<sub>600</sub> de 0,5 – 0,7. La inducción de la expresión se realizó con 1 mM IPTG, manteniendo el cultivo en las mismas condiciones durante 4 horas más.

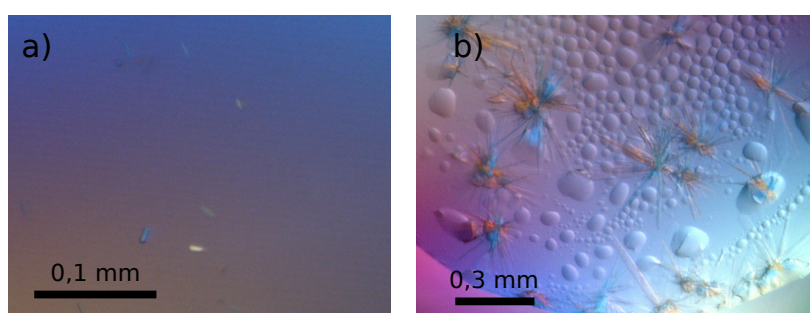
La purificación de DacB<sub>20–238</sub> se produjo mediante cromatografía de afinidad en columna His Trap<sup>TM</sup>HP Ni-NTA de 1 mL (GE Healthcare, UK), por llevar una etiqueta de 6 histidinas asociada a la secuencia. Esta etiqueta se eliminó mediante una incubación de 16 horas a 4 °C con la proteasa TEV, seguida de una purificación en columna His Trap<sup>TM</sup>HP (Saleh *et al.*, 2013). La DacB<sub>20–238</sub> final, sin la etiqueta de histidinas, se dializó frente a 20 mM Tris pH 8,0 y se concentró hasta alcanzar 10 mg/mL.

### 5.2.2. Cristalización y optimización

La cristalización de DacB<sub>20–238</sub> se llevó a cabo en varias fases. En primer lugar, se realizaron ensayos de cristalización mediante el método de difusión de vapor en gota sentada con técnicas de alto rendimiento (Sección 3.2.5), ensayando los paquetes comerciales Crystal Screen, Crystal Screen 2, INDEX, JCSG, PACT Suite y JBScreen Classic 1-4 (Tabla 3.1). Cada prueba de cristalización se realizó a 18 °C en una gota de 250 nL de proteína en solución y 250 nL de precipitante, todo ello equilibrado contra 65 µL de precipitante. Pasados 12 días se observó un único cristal de pequeño tamaño ( $\simeq$  0,02 mm) en una condición de cristalización que contenía 0,5 mM ZnCl<sub>2</sub>, 0,1 M Tris-HCl pH 7,5 y 16 % (p/v) PEG 6000. Se realizó un escalado de dicha condición, añadiendo ligeras variaciones de concentración y pH en los componentes de la condición de cristalización inicial. Se obtuvo un mayor número de cristales en 10 mM ZnCl<sub>2</sub>, 0,1 M Tris-HCl pH 8,0 y 20 % (p/v) PEG 6000, aunque mantenían la morfología y el tamaño del cristal original (Figura 5.3a). Ante esta situación, debido a la existencia de la región N-terminal presuntamente flexible (Ser20 – Ala55), se decidió caracterizar el efecto de distintos detergentes (Tabla 3.1) sobre la condición de cristalización previa, con la esperanza de estabilizar este fragmento y mejorar el empaquetamiento. De este modo, se obtuvieron cristales de mayor tamaño (0,05 a 0,4 mm) en presencia del detergente CYMAL-3, a una concentración de 35 mM (Figura

5.3b). No obstante, estos cristales eran extremadamente finos y su crecimiento superpuesto complicó la extracción de monocristales para los experimentos de difracción.

Un apunte interesante es el hecho de que la proteína cristalizara en presencia de  $\text{ZnCl}_2$ . Como se indicó anteriormente, a partir de la secuencia proteica se predijo la existencia de un sitio de unión a zinc en el dominio catalítico de DacB, por lo que el  $\text{ZnCl}_2$  podría estar contribuyendo a la estabilidad del sitio y, por tanto, a la estabilidad de la proteína en general. Para probar la importancia del zinc en la etapa de cristalización, se realizaron los mismos ensayos descritos anteriormente en presencia de distintas concentraciones de EDTA (de 20 mM a 0,2 M), obteniendo siempre agregados amorfos. De este modo se reforzó la idea de la existencia del sitio de unión a zinc en DacB y de la importancia del zinc en la estabilización de la proteína durante la etapa de cristalización.



**Figura 5.3: Cristales de DacB<sub>20–238</sub>.** a) Cristales con forma de prisma, de pequeño tamaño, en una gota de 2  $\mu\text{L}$ . b) Cristales de mayor tamaño, en una gota de 2  $\mu\text{L}$ . Se puede observar la existencia de un núcleo del que crecen láminas finas superpuestas.

Debido a la agrupación de los cristales en morfología de erizo, se intentó la cristalización en presencia de diferentes alcoholes (metanol, etanol e isopropanol) a diferentes concentraciones, ya que estos compuestos favorecen la precipitación y pueden alterar la dinámica del proceso (Yoshikawa *et al.*, 2012). También se intentaron diferentes técnicas de sembrado (Sección 3.2.5), con la intención de aumentar el grosor de los cristales. Finalmente, debido a su fragilidad, se realizó una deshidratación de algunos cristales aumentando la concentración de PEG desde el 20 % (p/v) hasta el 32 % (p/v), en saltos de 2 % cada 24 horas.

#### 5.2.3. Determinación estructural

Se recogieron cristales en todas las etapas de optimización descritas anteriormente y se difractaron en las líneas de radiación sincrotrón ID29 del ESRF, BL13 (Xaloc) de ALBA y X06DA (PXIII) del SLS. Todos los cristales fueron crioprottegidos en 20 % (v/v) glicerol.

La morfología de placa fina de estos cristales, junto con el hecho de que el grupo espacial parecía ser P2 o P2<sub>1</sub>, dificultó la tarea de recoger un conjunto de datos completo. Para que un conjunto de datos de un grupo espacial monoclinico sea completo, se requiere recoger la difracción de al menos 180° de rotación del cristal (Dauter, 1999). En el caso de DacB<sub>20–238</sub>, la delgadez de los cristales reducía drásticamente la difracción en los bordes, obteniéndose una completitud habitualmente no superior al 70 %. Otro problema común en los cristales de DacB<sub>20–238</sub> era la elevada anisotropía del patrón de difracción. Si bien la resolución global alcanzaba valores de

entre 3 y 3,5 Å, la resolución en el eje corto no solía superar los 8 Å, en el mejor de los casos, lo que reduce significativamente la calidad de los datos.

En la Tabla 5.1 se recogen las estadísticas de difracción del mejor cristal de DacB<sub>20–238</sub> cuya completitud para la última capa de resolución alcanzó valores óptimos. No obstante, la resolución en el eje corto no superaba los 7,5 Å, pese a que globalmente se alcanzase una resolución de 3 Å. Indexado e integrado se realizaron con XDS e iMosflm, mientras que el escalado se realizó con XSCALE y Aimless respectivamente.

**Tabla 5.1:** Datos y estadísticas de difracción de DacB<sub>20–238</sub>.

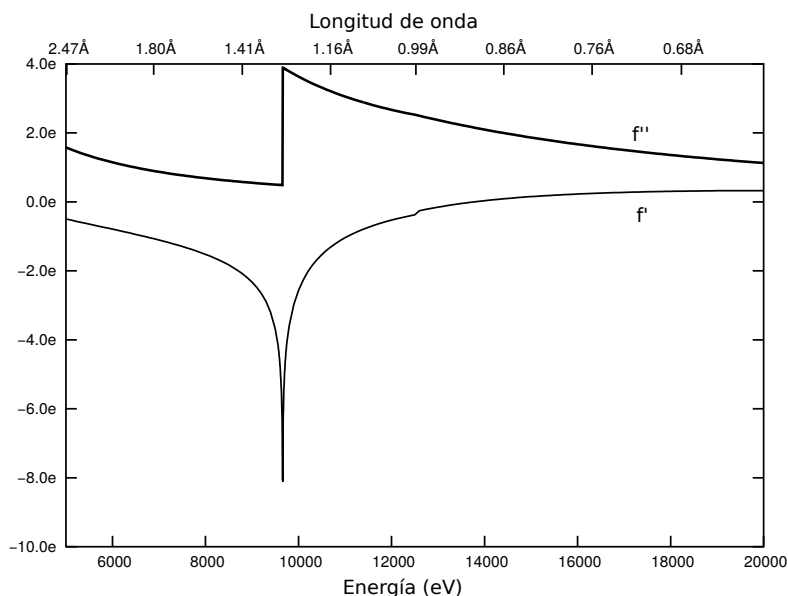
PARÁMETRO	DACB <sub>20–238</sub>
Grupo espacial	P2 <sub>1</sub>
$a, b, c$ (Å)	103,3, 37,7, 106,3
$\alpha = \beta = \gamma$ (°)	90,0, 109,8, 90,0
Resolución (Å)	48,6 – 3,0 (3,16 – 3,00)
Longitud de onda (Å)	1,2812
Línea	ID29 (ESRF)
Temperatura (K)	100
Nº de reflexiones	
Totales	53269
Únicas	15936
$\langle I/\sigma I \rangle$	6,4 (2,0)
$R_{\text{pim}}$	0,11 (0,41)
$CC_{1/2}$	0,98 (0,78)
Completitud (%)	99,6 (99,7)
Multiplicidad	3,3 (3,4)

Los valores entre paréntesis hacen referencia a la última capa de resolución.

Según las predicciones realizadas en base al cálculo del coeficiente de Matthews, cabe la posibilidad de que existan 3 ó 4 moléculas en la unidad asimétrica (para 3 moléculas, coeficiente de Matthews = 2,61 Å<sup>3</sup>/Da, contenido de solvente = 52,9 %, probabilidad = 0,65; para 4 moléculas, coeficiente de Matthews = 1,96 Å<sup>3</sup>/Da, contenido de solvente = 37,2 %, probabilidad = 0,29). Lo más probable es que existan 3 moléculas por unidad asimétrica, no sólo por los resultados derivados del coeficiente de Matthews sino también por la baja resolución de los datos (3 Å), sinónimo de la existencia de un empaquetamiento débil y un alto porcentaje de solvente.

En el momento en el que obtuvieron los primeros datos, no existía un modelo de reemplazo adecuado para intentar un faseado mediante esta estrategia. Así pues, se intentó obtener un modelo parcial mediante la fusión de fragmentos de otras proteínas, que cubriese parte de la secuencia y permitiese determinar el plegamiento general de DacB<sub>20–238</sub>. Para ello se construyeron modelos tanto de forma manual como con programas como Sculptor, sin conseguir soluciones de reemplazo fiables. También se utilizó Balbes, que realiza búsquedas exhaustivas de fragmentos similares y los combina para, posteriormente, realizar varias estrategias de reemplazo molecular. Esta aproximación tampoco devolvió resultados fiables.

Ante esta situación y ante la posible existencia de un sitio de unión a zinc, los datos de la Tabla 5.1 se recogieron a la longitud de onda coincidente con el pico de la banda K de absorción de dicho elemento (Figura 5.4), con la intención de fasear por SAD (Sección 3.3.3.4). El espectro de fluorescencia experimental determinaba la presencia de zinc en el cristal pero, debido a la utilización de  $\text{ZnCl}_2$  en la condición de cristalización, no podía demostrarse la existencia del sitio específico de unión a zinc ni que los átomos de zinc estuvieran tan ordenados como para generar suficiente dispersión anómala aprovechable.



**Figura 5.4: Espectro de fluorescencia teórico del zinc.** Se observa que el pico de máxima dispersión (banda K de absorción) se sitúa entre 1,2 y 1,3 Å.

En la Tabla 5.2 se muestran las estadísticas relacionadas con el faseado mediante Zn-SAD de  $\text{DacB}_{20-238}$ . La señal anómala obtenida apenas superaba los 6 Å de resolución. La búsqueda de átomos pesados encontró un único sitio con ocupación 1,0. En general, las estadísticas del faseado se encuentran lejos de los valores adecuados para una correcta resolución por SAD. Debido a ello, fue imposible determinar la subestructura correcta y, por tanto, unas fases preliminares. Además, sigue sin poder determinarse la existencia del sitio de unión a zinc, ya que si bien se predice un sitio de zinc con ocupación del 100 %, la señal anómala podría generarse a partir de algún átomo de  $\text{Zn}^{2+}$  de la condición de cristalización, estabilizado mediante unión inespecífica a alguna cadena lateral expuesta.

Posteriormente, el Centro de Genómica Estructural de Enfermedades Infecciosas (CSGID, NIH, EEUU) resolvió la proteína VanXYg de *Enterococcus faecalis* (código PDB 4F78, Meziane-Cherif *et al.*, 2014). Esta proteína presentaba una similitud del 39,7 % y una identidad del 27,0 % con el dominio catalítico teórico de DacB (Ala55 – Asp238). Además, contenía un sitio de unión a zinc formado por dos histidinas y un aspártico, conservadas en DacB (His153, Asp160 e His207). La existencia de esta estructura permitió realizar tanto reemplazos moleculares como MR-SAD, es decir, una búsqueda mediante reemplazo molecular acoplada a un faseado mediante la señal anómala del zinc. Esta estrategia se realiza con la finalidad de potenciar dos métodos que, en este caso, por separado eran débiles. De este modo, se obtuvieron unos resultados preliminares que

**Tabla 5.2:** Estadísticas relacionadas con el faseado experimental de DacB<sub>20–238</sub>.

PARÁMETRO	Zn
Longitud de onda (Å)	1,2812
Resolución (Å)*	5,5
$CC_{\text{anom}}$	0,52
$CC_{\text{all}}/CC_{\text{weak}}$	0,22 / 0,11
CFOM	0,33
Sitios**	1
Contraste directa (%)	65,9
Contraste inversa (%)	66,4
$CC_{\text{map}}$ directa ***	0,87
$CC_{\text{map}}$ inversa ***	0,88

\* El límite de resolución de la señal anómala se establece para  $\langle d''/\text{sig} \rangle = 1,3$ .

\*\* Se consideran los que presentan una ocupación superior al 70 %.

\*\*\* Calculado para la máxima resolución del conjunto de datos.

sirvieron para determinar la ubicación de algunos elementos aislados de estructura secundaria, además de aportar evidencias sólidas sobre la existencia del sitio de unión a zinc en DacB. Por tanto, el modelo estaba lejos de ser fiable y completo, aunque parecía confirmar la existencia del sitio de unión a zinc.

Tras numerosos intentos con varios programas y estrategias diferentes, el mejor resultado se obtuvo con MrBUMP. La ventaja de este programa en este caso fue la capacidad de probar varios modelos y estrategias combinadas por fuerza bruta, generando cientos de resultados que se ordenan en base a los factores  $R$  de refinamiento tras aplicar cada estrategia. En concreto, MrBUMP ensayó 9 modelos diferentes, todos ellos basados en la VanXYg de *E. faecalis*. La mejor solución la encontró con Phaser sobre un modelo truncado de VanXYg cuyas cadenas laterales se habían sustituido, mediante Chainsaw, por las cadenas de DacB<sub>20–238</sub>. Posteriormente, realizó varios ciclos recursivos de construcción y refinamiento con Buccaneer y Refmac5, devolviendo un modelo con dos moléculas por unidad asimétrica cuyos factores  $R$  se situaban en  $R_{\text{work}} = 0,43$  y  $R_{\text{free}} = 0,50$ .

Finalmente, tras revisar, corregir y reconstruir manualmente este modelo en Coot, se obtuvo la mejor estructura posible para los cristales de DacB<sub>20–238</sub>. Las estadísticas de refinamiento, realizado entre Phenix y Refmac5, se muestran en la Tabla 5.3. Como se puede comprobar, la geometría del modelo está muy alejada de los valores óptimos, hecho que se relaciona directamente con la baja resolución, la elevada anisotropía y la falta de completitud en las últimas capas de resolución.

#### 5.2.4. Análisis estructural

Los datos obtenidos con MrBUMP y optimizados de forma manual permitieron trazar 2 moléculas de las 3 ó 4 posibles por unidad asimétrica, con longitudes de 196 y 184 residuos respectivamente. La primera cadena abarca desde la posición 43 hasta el Asp238, si bien los primeros 12 aminoácidos (desde la posición 43 a la 54) se han sustituido por alaninas debido a la baja calidad de la densidad electrónica de esa zona. La segunda cadena abarca desde la Ala55 hasta el Asp238.

**Tabla 5.3:** Estadísticas de refinamiento de DacB<sub>20–238</sub>.

PARÁMETRO	DACB <sub>20–238</sub>
Resolución (Å)	48,6 – 3,0
$R_{\text{work}}/R_{\text{free}}$	0,39 / 0,47
Nº de átomos	
Proteína (no H)	3019
Zinc	2
Agua	0
Factores B (Å <sup>2</sup> )	
Proteína	44,6
Zinc	44,8
RMSD de enlaces	
Distancias (Å)	0,016
Ángulos (°)	2,360
Ramachandran (%)	
Favorables	48,1
Desfavorables	24,2

#### 5.2.4.1. Estructura general y sitio de unión a zinc

La estructura de DacB<sub>20–238</sub> muestra un dominio globular (Ala55 a Asp238), correspondiente al dominio catalítico de DacB, acompañado en el extremo N-terminal de un fragmento flexible y desordenado, cuya determinación estructural no ha sido posible (Figura 5.5). Si bien no se puede determinar con precisión la posición de la mayoría de los residuos, se distinguen ciertos elementos de estructura secundaria dentro del dominio catalítico. De esta forma, se puede determinar que el plegamiento general corresponde a una serie de hebras  $\beta$  envueltas por al menos dos grupos de hélices  $\alpha$ .

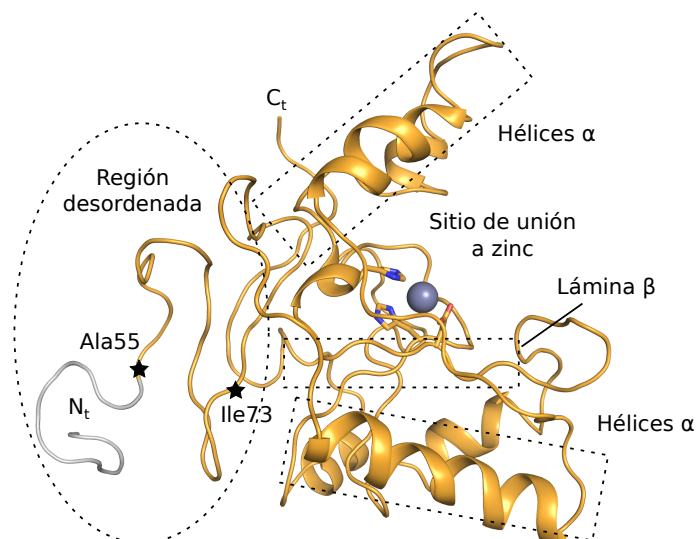
El sitio de unión a zinc, formado por los residuos His153, Asp160 e His207, se encuentra en el centro de la estructura globular, en una cavidad expuesta al solvente, sobre lo que podrían ser unas hebras  $\beta$  antiparalelas. La baja calidad tanto de los parámetros geométricos globales como de la densidad electrónica a lo largo de todo el modelo impide realizar una descripción estructural más precisa.

#### 5.2.4.2. Empaquetamiento cristalino

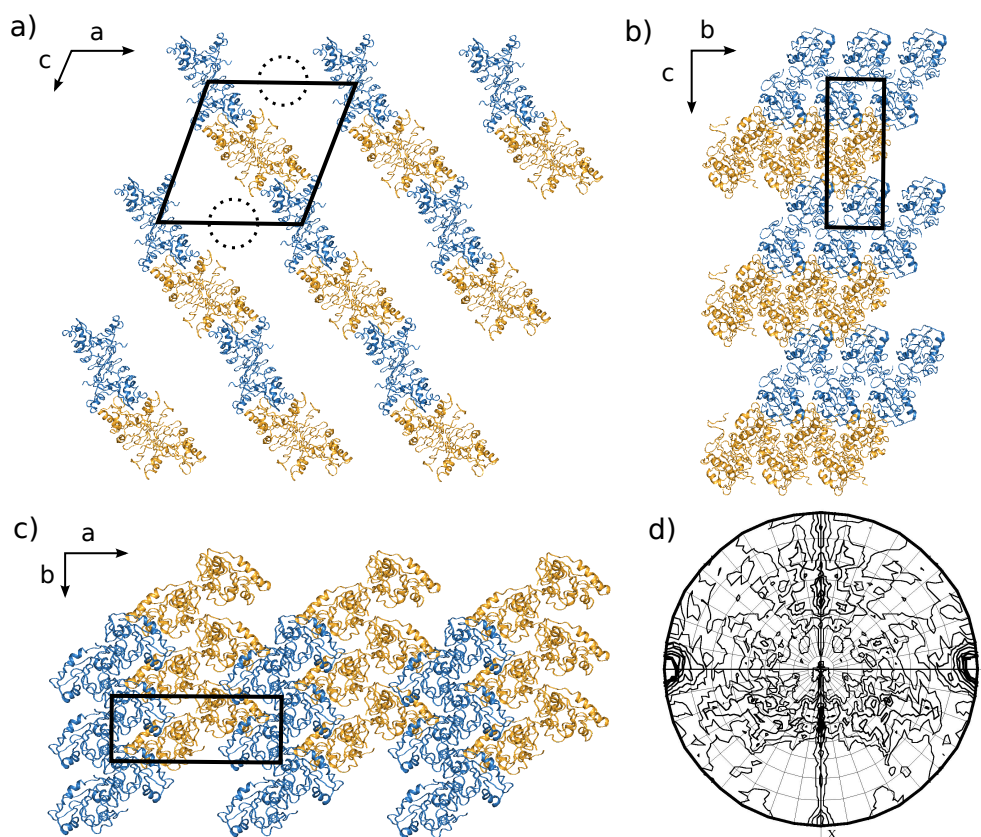
Como se indicó anteriormente, el grupo espacial en el que cristalizó DacB<sub>20–238</sub> es el monoclinico P2<sub>1</sub>, siendo el eje  $b$ , por convenio, el binario helicoidal. En este caso,  $a \simeq c \simeq 104$  Å, mientras que el eje  $b$  es mucho más corto (37,7 Å). Si se sigue la predicción obtenida mediante el coeficiente de Matthews, las 3 moléculas por unidad asimétrica darían lugar a 6 moléculas en la celdilla unidad, algo que parece bastante plausible a la vista el espacio libre que queda en la celdilla (Figura 5.6).

Al analizar el empaquetamiento general se observa claramente que las dos moléculas que se han encontrado forman capas paralelas separadas entre sí por el hueco que debería ocupar la tercera cadena (5.6). Por tanto, el papel de esta tercera molécula es crucial para la estabilidad general del cristal. El hecho de que no se pueda determinar su ubicación responde a una supuesta





**Figura 5.5: Estructura general de DacB<sub>20-238</sub>.** Se observa un plegamiento globular formado por una posible lámina  $\beta$  rodeada de dos zonas de hélices  $\alpha$ , englobando al sitio de unión a zinc (representado con varillas y esfera). La región N-terminal (Ser20 – Ile73) aparece bastante desordenada, lo que impide construir la secuencia hasta la posición 43 ( $N_t$ ). Desde el residuo 43 hasta el 54 (color blanco) se ha trazado una cadena de poli-Ala, debido a la incapacidad de determinar las posiciones aproximadas de las cadenas laterales. A partir de la Ala50 se han podido construir las cadenas laterales aproximadas, aunque el desorden continúa hasta la Ile73.



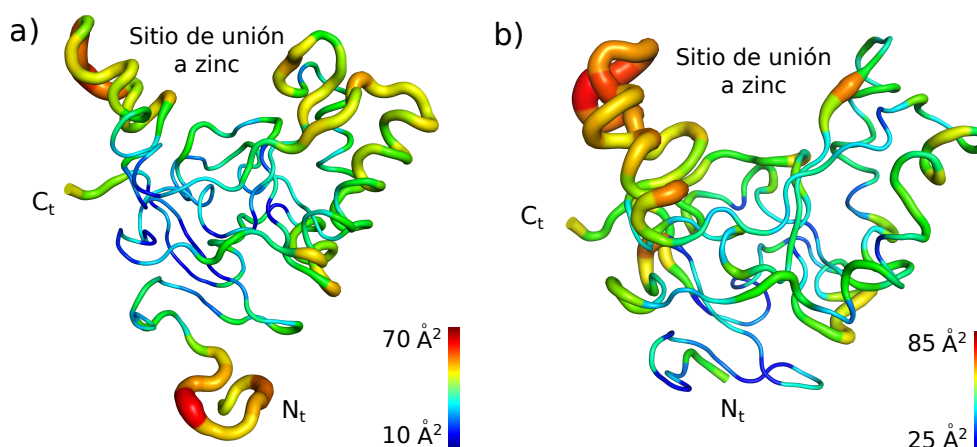
**Figura 5.6: Simetría y empaquetamiento de DacB<sub>20-238</sub>.** Las dos cadenas construidas se representan en naranja y azul. El recuadro negro hace referencia a la celdilla unidad, indicándose además la orientación de cada eje en cada caso. **a)** Plano ( $a, c$ ). Se observa la asociación en capas paralelas de las dos moléculas construidas, mientras que la tercera molécula completaría los contactos cristalinos (círculo punteado). **b)** Plano ( $b, c$ ). **c)** Plano ( $a, b$ ). Función de autorrotación calculada para  $\chi = 180^\circ$  mediante Molrep, mostrando los picos característicos de un grupo puntual  $2/m$ .

inestabilidad de los contactos con las moléculas vecinas, lo que se traduce en un cristal frágil y de baja calidad, correlacionándose con los resultados observados durante la cristalización de DacB<sub>20–238</sub>.

Las dos moléculas construidas se relacionan entre sí enfrentando los sitios de unión a zinc, quedando las regiones N-terminales orientadas hacia el sentido opuesto, enfrentadas con las regiones N-terminales de sus mismas cadenas vecinas. La imposibilidad de construir la región N-terminal al completo impide determinar hacia qué zona del cristal se orienta, aunque únicamente hay espacio para su localización en la zona entre ambas capas. De esta forma, el espacio actualmente vacío se vería ocupado por la tercera molécula y los extremos N-terminales de las 3 cadenas.

#### 5.2.4.3. Agitación térmica

En general, las dos moléculas muestran unos factores B medios óptimos para los datos de los que se dispone (cadena A = 41,5 Å<sup>2</sup>, cadena B = 48,0 Å<sup>2</sup>). No obstante, existe un lazo (Asp71 a Ala80), expuesto al solvente, que presenta unos factores B sensiblemente superiores al resto de la estructura, con valores que alcanzan los 80 Å<sup>2</sup> (Figura 5.7).



**Figura 5.7: Representación de factores de agitación térmica de DacB<sub>20–238</sub>.** Los valores cercanos al azul reflejan baja agitación térmica, mientras que un virado hacia el rojo implica mayor agitación térmica y mayor flexibilidad. Para mayor claridad, se han indicado la posición del sitio de unión a zinc y los extremos N y C-terminales (N<sub>t</sub> y C<sub>t</sub> respectivamente). **a)** Cadena A. La región de poli-Ala muestra una agitación significativamente superior al resto de la estructura. **b)** Cadena B.

#### 5.2.4.4. Otras moléculas y solvente

En los sitios de unión a zinc se encontraron densidades débiles correspondientes probablemente a Zn<sup>2+</sup> coordinado, por lo que se han modelado en la estructura final. Sus factores de agitación térmica son muy similares a los de las cadenas laterales que les rodean, aunque este hecho no es especialmente relevante ya que la baja calidad de los datos tiende a carecer de precisión en distancias cortas. Debido también a la baja calidad de los datos, no se encontraron ni modelaron moléculas de agua.

### 5.3. Segunda construcción (DacB<sub>61–238</sub>)

Tras la imposibilidad de obtener un modelo fiable y completo de DacB<sub>20–238</sub>, se decidió resolver una construcción que abarcase únicamente el dominio catalítico de DacB, dejando fuera tanto el péptido señal como la región flexible. El final de la construcción se mantuvo en el último residuo de la DacB, el Asp238, mientras que el inicio se situó en la Gly61, lo que supone un total de 178 aminoácidos, un peso molecular teórico de 20,1 kDa y un pI teórico de 4,8 (Figura 5.8).

10	20	30	40	50	60
MKKRYLVLTA	LLALSAAACS	QEKTKNEDGE	TKTEQTAKAD	GTVGSKSQGA	AQKKA <del>EVVNK</del>
70	80	90	100	110	120
GDYYSIQGKY	DEIIVANKHY	PLSKDYNPGE	NPTAKAELVK	LIKAMQEAGF	PISDHYSGFR
130	140	150	160	170	180
SYETQTKLYQ	DYVNQDGKAA	ADRYSARPGY	SEHQTGLAFD	VIGTDGDLVT	EEKAAQWLLD
190	200	210	220	230	
HAADYGFVVR	YLGKKEKETG	YMAEEWHLRY	VGKEAKEIAA	SGLSLEEYYG	FEGGDYVD

**Figura 5.8: Secuencia de la construcción DacB<sub>61–238</sub>.** En gris se muestra el péptido señal, mientras que en rojo se muestra el dominio catalítico. El sombreado amarillo se corresponde con la secuencia de la construcción de DacB<sub>61–238</sub>.

El principal motivo para determinar el primer residuo de la construcción del dominio catalítico fue la similitud existente entre la predicción de la estructura secundaria de DacB y el dominio N-terminal de VanXYg de *E. faecalis* (código PDB 4F78, Meziane-Cherif *et al.*, 2014), que, como se ha dicho anteriormente, era la única estructura similar resuelta hasta la fecha. Era de esperar que el dominio catalítico de DacB solapase en alto grado con el mismo dominio de VanXYg. Así pues, se eligió como primer aminoácido de la construcción la Gly61, por ser coincidente con el primer residuo visible de VanXYg (Figura 5.9).

#### 5.3.1. Expresión y purificación

La obtención de DacB<sub>61–238</sub> fue realizada por el grupo del Prof. Sven Hammerschmidt, del Departamento de Genética Microbiana de la Universidad de Greifswald, Alemania, siguiendo el protocolo descrito para DacB<sub>20–238</sub> (Sección 5.2.1). La proteína final se obtuvo a una concentración de 10 mg/mL en 20 mM Tris-HCl pH 8,5.

#### 5.3.2. Cristalización y resultados preliminares

Se iniciaron los experimentos de cristalización de DacB<sub>61–238</sub> ensayando variaciones de la condición de cristalización de DacB<sub>20–238</sub> (0 – 15 mM ZnCl<sub>2</sub>, 0,1 M Tris-HCl pH 6,0 – 10,0, 14 – 24 % (p/v) PEG 6000), en presencia y ausencia de CYMAL-3, sin conseguir resultados positivos. En paralelo se realizaron los ensayos de alto rendimiento de la misma forma que en el caso de DacB<sub>20–238</sub> (Sección 5.2.2), apareciendo cristales con una morfología novedosa y de elevada calidad óptica en varias condiciones nuevas y similares. Las mejores condiciones, basadas en 0,2 M CaCl<sub>2</sub>/LiCl, 0,1 M Tris-HCl pH 7,5 – 8,5 y diversas concentraciones de PEG 4000,

Péptido señal			
DacB	1	MKKRYLVLTALLALSLAAC <b>SQ</b> EKTKNEDGETKTEQTAKADGTVGSKSQGA	50
VanXYg	1	-----	0
DacB	51	AQKKA <b>EVVNKGDYYSIQGKYDEI</b> -----IVANKHYPLSKDYN-----PGE	90
		...: :..:    :.: :    : .    .:	
VanXYg	1	-----GMMKTIELEKEEIYCGNLLLVNKNYPL-RDNNVKGLVPAD	39
DacB	91	-----NPTAKAELVKLIKAMQE-----AG-----FPISDHYSGFRSYETQTKLYQ	130
		... :..:..:..:      . :    :   .  .  .:	
VanXYg	40	IRFPNILMKRDVANVLQLIFEKISAGNSIVPV---SGYRSLEEQTAIYD	85
DacB	131	DYVNQDGGKAAADRY <b>SARPGYSEHQ</b> TGLAFDVIG---TDGDLVTEEKAAQW	177
		..:..: :..:..: .  .  .:           .   :  . .  .:..:..:..:	
VanXYg	86	GSLKDNGEDFTRKYVALPNHSEHQ <b>TGLA</b> ID-LGLNKKDIDFIRPDFPYDG	134
DacB	178	LLDH-----AADYGFVVRYLKGKEKETGYMAE <b>EWHLRYVG-KEAKEIAASG</b>	222
		: .  .        .    .  .:   .  .  .        .  .: .:..:	
VanXYg	135	ICDEFRAAPDYGFTQRYARDKEEITGISHEPW <del>HF</del> RYVGYPH <b>SKIMQ</b> ENG	184
DacB	223	LSLEYYGF-----EGGDYVD-----	238
		.        .  .    .: .:	
VanXYg	185	FSLEEY <b>TQFIKAYLEDNKYL</b> FEQA <b>HRAE</b> IEIYYVPAKDDKTLIKIPENCV	234
DacB	239	-----	238
VanXYg	235	YQISGNNIDGFVVTIWRKTDD	255

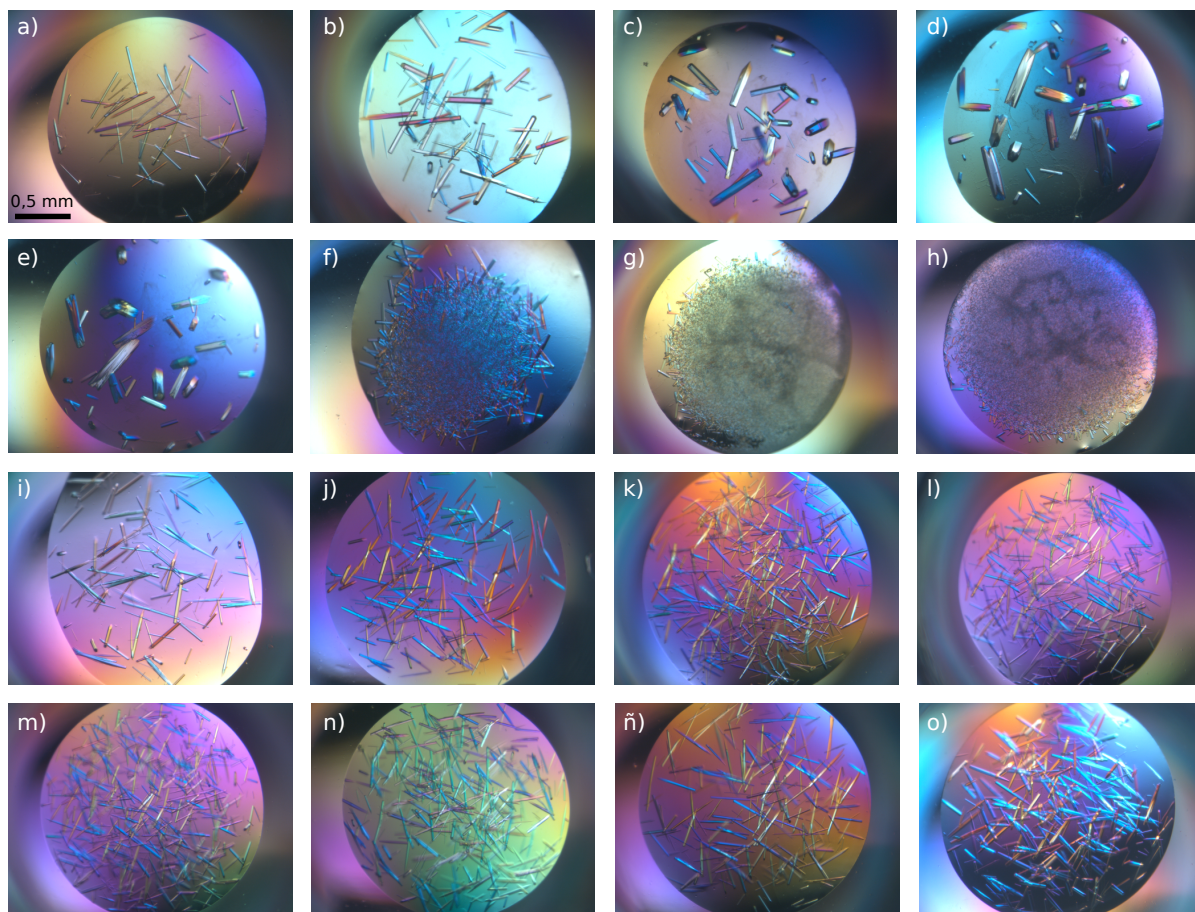
**Figura 5.9: Comparación de la secuencia de DacB y VanXYg.** En la secuencia de DacB, se representa en rojo los aminoácidos correspondientes al dominio catalítico. El recuadro azul se corresponde al solapamiento entre el dominio catalítico de DacB y VanXYg. Se puede observar que el primer aminoácido de VanXYg es la Gly1, correspondiente a la Gly61 en DacB.

fueron optimizadas. Se pudo observar que el comportamiento del crecimiento de DacB<sub>61–238</sub> era especialmente sensible a la concentración de PEG 4000, ya que pequeñas variaciones de este componente provocaban cambios progresivos pero radicales en la morfología de los cristales (Figuras 5.10a a 5.10h). Los mejores cristales se obtuvieron en 0,2 M CaCl<sub>2</sub>, 0,1 M Tris-HCl pH 8,5 y 18 % (p/v) PEG 4000 (Figura 5.10b). Además, el crecimiento de todos los cristales era extremadamente rápido (entre 12 y 48 horas), algo poco habitual y que indica una gran afinidad de las moléculas de DacB<sub>61–238</sub> entre sí. Los cristales crecidos en 0,2 M de LiCl, 0,1 M Tris-HCl pH 8,5 y concentraciones variables de PEG 4000 mostraban la misma morfología que al usar CaCl<sub>2</sub>, pero las alteraciones frente a la variación de PEG 4000 eran menos drásticas (Figuras 5.10i a 5.10o).

Algunos de estos cristales se crioprotegieron en 10 % (v/v) glicerol y se midieron en el difractómetro del Instituto de Química-Física Rocasolano (CSIC, Madrid; Sección 3.3.1.1) alcanzando una resolución de 2,5 Å (Tabla 5.4, DacB<sub>61–238</sub>(1)). No obstante, la determinación de la celdilla unidad parecía ser imprecisa ya que, según la predicción realizada mediante el coeficiente de Matthews, el volumen de la celdilla unidad era inferior al mínimo necesario para albergar al menos una molécula de DacB<sub>61–238</sub>.

Con la intención de mejorar la precisión de los datos y descartar limitaciones en la medida debidos a la tecnología del difractómetro, se recogieron datos de difracción de numerosos cristales





**Figura 5.10: Cristales de DacB<sub>61-238</sub> en gradiente de PEG 4000.** a – h) Cristales en 0,2 M CaCl<sub>2</sub>, 0,1 M Tris-HCl pH 8,5 y concentración variable de PEG 4000. El panel a muestra los cristales en 16 % (p/v) PEG 4000, aumentando un 2 % en cada panel, terminando con un 30 % (p/v) PEG 4000 en el panel h. A concentraciones altas de PEG 4000 (26 – 30 %) la nucleación es masiva y los cristales extremadamente pequeños. A concentraciones bajas (16 – 18 %) la nucleación es mucho menor, lo que genera cristales de mayor longitud. A concentraciones intermedias (20 – 24 %), pese a que la nucleación es baja, los cristales crecen demasiado, acumulando defectos e invaginaciones en los extremos. Con concentraciones menores del 14 % no se forman cristales (no mostrado). i – o) Cristales en 0,2 M LiCl, 0,1 M Tris-HCl pH 8,5 y concentración variable de PEG 4000. El panel i muestra los cristales en 16 % (p/v) PEG 4000, aumentando un 2 % en cada panel, terminando con un 30 % (p/v) PEG 4000 en el panel o. En contraposición a lo que sucede en presencia de CaCl<sub>2</sub>, el aumento en la concentración de PEG 4000 no provoca diferencias en el tamaño de los cristales, sino únicamente un aumento en la tasa de nucleación, lo que conlleva la ausencia de defectos por crecimiento excesivo. Todos los paneles están a la misma escala que el panel a.

de diferentes condiciones y morfologías en las líneas BL13 (Xaloc) de ALBA y X06DA (PXIII) del SLS, todos crioprottegidos en 10 % (v/v) glicerol. Los datos obtenidos son de muy alta calidad, alcanzando resoluciones de hasta 1,6 Å y valores estadísticos habituales de patrones de esta resolución. No obstante, el procesamiento de cristales aparentemente iguales mostraba celdillas y grupos espaciales distintos. En la Tabla 5.4 se muestran las estadísticas de dos cristales recogidos en líneas de sincrotrón, aparentemente iguales (DacB<sub>61-238</sub>(2) y DacB<sub>61-238</sub>(3)).

En general, conjuntos de datos de cristales aparentemente iguales mostraban celdillas unidad de dos tamaños diferentes (siendo la única diferencia  $a = b \simeq 37$  Å o  $a = b \simeq 64$  Å), con grupos espaciales que oscilaban entre P3, P321, P312, P3<sub>1</sub>12, P622 o P6<sub>2</sub>22. Además, era habitual la aparición de errores en el indexado de la celdilla durante el procesamiento, así como la existencia de diferentes subconjuntos de reflexiones dentro de un mismo conjunto de datos que satisfacían

### III. RESULTADOS Y DISCUSIÓN

**Tabla 5.4:** Datos y estadísticas de difracción preliminares de DacB<sub>61–238</sub> correspondientes a tres cristales aparentemente iguales (DacB<sub>61–238</sub>(1), DacB<sub>61–238</sub>(2) y DacB<sub>61–238</sub>(3)).

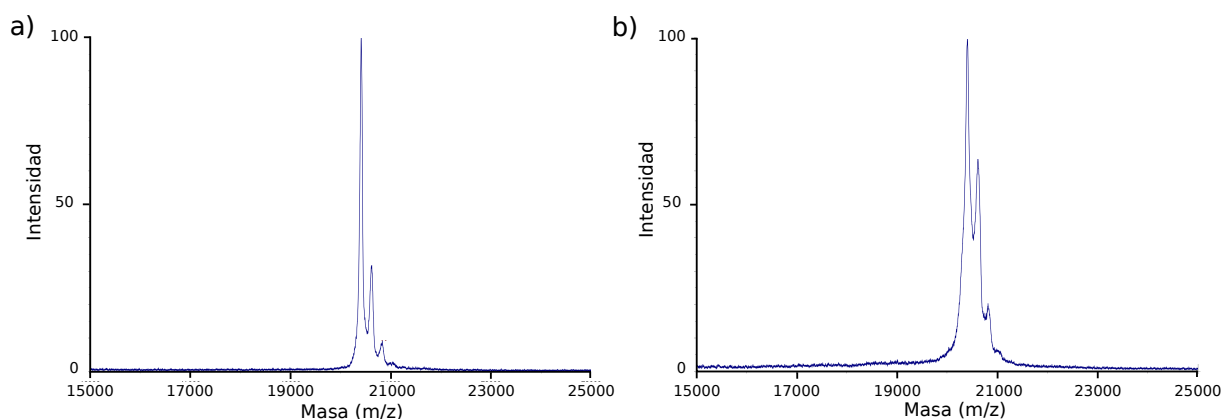
PARÁMETRO	DACB <sub>61–238</sub> (1)	DACB <sub>61–238</sub> (2)	DACB <sub>61–238</sub> (3)
Grupo espacial	P321	P6 <sub>2</sub> 22	P3 <sub>1</sub> 12
$a, b, c$ (Å)	37,04, 37,04 111,70	37,69, 37,69, 110,2	64,37, 64,37, 111,53
$\alpha = \beta = \gamma$ (°)	90,0, 90,0, 120,0	90,0, 90,0, 120,0	90,0, 90,0, 120,0
Resolución (Å)	18,6 – 2,5 (2,60 – 2,50)	36,7 – 2,0 (2,06 – 2,00)	39,4 – 1,7 (1,73 – 1,70)
Longitud de onda (Å)	1,54180	1.00000	0.97949
Línea	Ánodo rotatorio (IQFR)	PXIII (SLS)	BL13 (ALBA)
Temperatura (K)	100	100	100
Nº de reflexiones			
Totales	10531	108561	565834
Únicas	3322	3534	29393
$\langle I/\sigma I \rangle$	5,0 (2,1)	9,8 (2,2)	15,5 (3,1)
$R_{\text{pim}}$	0,118 (0,335)	0,094 (0,371)	0,029 (0,242)
$CC_{1/2}$	0,98 (0,76)	0,99 (0,68)	0,99 (0,92)
Compleitud (%)	98,4 (97,4)	99,7 (98,6)	99,9 (99,7)
Multiplicidad	3,2 (3,4)	30,7 (31,0)	19,3 (19,9)

Los valores entre paréntesis hacen referencia a la última capa de resolución.

diferentes geometrías de celdilla. Todos los conjuntos de datos se indexaron e integraron tanto con XDS como con iMosflm, siendo escalados con XSCALE o Aimless respectivamente.

Al realizar la predicción del contenido de la celdilla mediante el coeficiente de Matthews, los resultados oscilaban entre 0, 1 y 2 moléculas por unidad asimétrica dependiendo del tamaño de la celdilla y del grupo espacial. En los casos en los que se predecían 1 ó 2 moléculas por unidad asimétrica, los reemplazos moleculares con el modelo parcial de DacB<sub>20–238</sub> no devolvían ninguna orientación fiable o incluso terminaban sin solución (tanto con Molrep como con Phaser). Otras aproximaciones realizadas con Balbes o MrBUMP tampoco dieron resultado.

Ante la posibilidad de una proteólisis de la muestra durante la cristalización, o de la cristalización de una proteína diferente de DacB<sub>61–238</sub>, contaminante de la purificación, se realizó una espectrometría de masas tanto de la proteína soluble como del contenido de varios cristales. Esta prueba fue realizada por el Servicio de Espectrometría de Masas del Instituto de Química-Física “Rocasolano” (IQFR, CSIC), empleando un espectrómetro Voyager-DE PRO (Applied Biosystems, EEUU) con un láser de nitrógeno ( $\lambda = 337$  nm, anchura de pulso 10 ns y frecuencia de 3 Hz) y una fuente iónica con extracción retardada. La ionización de las muestras se llevó a cabo por MALDI (*Matrix-Assisted Laser Desorption/Ionization*, ionización/desorción por láser asistida por matriz). Para ello, se utilizaron 15  $\mu\text{L}$  de cada muestra a una concentración aproximada de 10  $\mu\text{M}$ . Los resultados mostraron un único pico correspondiente a la masa teórica de DacB<sub>61–238</sub> (Figura 5.11), por lo que los cristales contenían la construcción original completa. Por tanto, la hipótesis más plausible para explicar la gran variedad de geometrías observadas era la existencia de una macla que alterase la determinación de los parámetros de celdilla y, consecuentemente, todas las estadísticas posteriores. Además, la macla también justificaría los problemas encontrados durante la realización del reemplazo molecular en los diferentes conjuntos de datos.



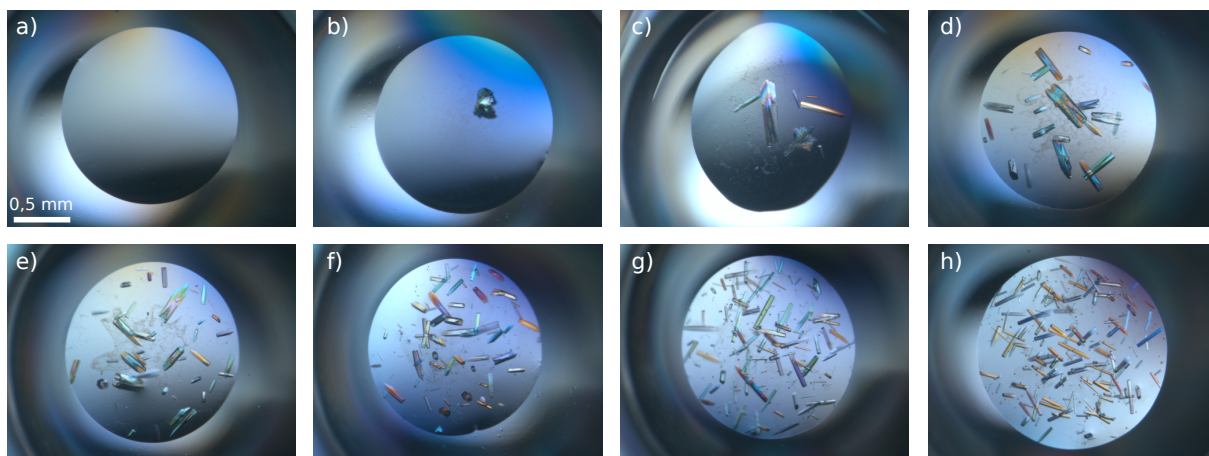
**Figura 5.11: Espectrometría de masas de DacB<sub>61–238</sub>.** a) Muestra de proteína soluble. b) Muestra del contenido de los cristales. En ambos casos se observa un peso molecular cercano a los 20 kDa, sin hallarse presencia de otra especie de menor tamaño.

### 5.3.3. Variantes de la cristalización para evitar maclas

Para intentar evitar la aparición de maclas en el proceso de cristalización de DacB<sub>61–238</sub>, se llevaron a cabo diferentes aproximaciones encaminadas a alterar la dinámica de agregación y el empaquetamiento. A continuación se enumeran, justificando el motivo de su uso:

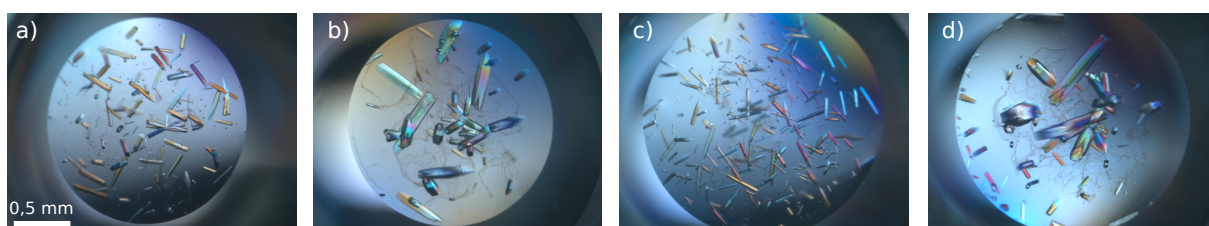
- **Variación de la concentración de todos los componentes de la condición de cristalización.** Como se ha observado con anterioridad, pequeñas variaciones en el porcentaje de PEG 4000 conllevan cambios sustanciales en la morfología de los cristales crecidos bajo CaCl<sub>2</sub>. En cambio, los cristales crecidos bajo LiCl sufrían menor variabilidad. Así pues, se llevó a cabo una variación progresiva de las concentraciones de todos los componentes de la condición de cristalización, tanto de forma individual como combinada. Los cambios más significativos se observaron al realizar un gradiente de CaCl<sub>2</sub>, observándose que a baja concentración se formaban pocos cristales pero de gran tamaño, que acumulaban numerosos defectos. A medida que se aumenta la concentración de CaCl<sub>2</sub> aumenta la nucleación, los cristales son más pequeños y se generan menos defectos, aunque mantienen la morfología original (Figura 5.12).
- **Variación de pH junto con la variación de concentración de los componentes anteriores.** Como se introdujo en la Sección 3.2.2, diferentes valores de pH alteran la solubilidad de la proteína para una misma condición de cristalización. Así pues, se realizaron ensayos de cristalización a diferentes puntos de pH (desde pH 4 a pH 9), al mismo tiempo que se probaban variaciones de concentración del resto de componentes de la condición de cristalización. Únicamente se obtuvieron cristales entre pH 7,5 y pH 9, con la morfología habitual.
- **Cristalización a 4 °C.** Con la finalidad de modificar la estabilidad de la proteína y ralentizar el crecimiento, intentando así evitar la formación de varias orientaciones en el mismo cristal, se realizaron experimentos de cristalización a 4 °C con las variaciones habituales de CaCl<sub>2</sub>, PEG 4000 y pH. En ningún caso se obtuvieron cristales.





**Figura 5.12: Cristales de DacB<sub>61-238</sub> en gradiente de CaCl<sub>2</sub>.** Experimentos de cristalización de DacB<sub>61-238</sub> en 0,1 M Tris-HCl pH 8,5, 22 % (p/v) PEG 4000 y concentración variable CaCl<sub>2</sub>. Los paneles a – h se corresponden con 0, 10, 25, 50, 100, 300, 400 y 460 mM CaCl<sub>2</sub> respectivamente. Todos los paneles están a la misma escala que el panel a.

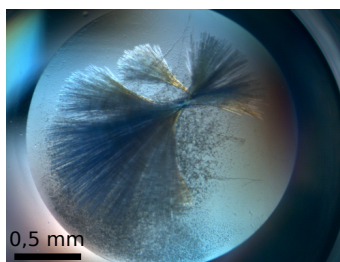
- **Cristalización bajo aceite.** Otra forma de reducir la tasa de nucleación y crecimiento, además de bajar la temperatura, es utilizar técnicas de difusión de vapor sin aislamiento (Sección 3.2.4.2). Para ello se enfrentó la proteína a la condición de cristalización habitual (0,2 mM CaCl<sub>2</sub>, 0,1 M Tris-HCl pH 8,5, 18 % (p/v) PEG 4000), cubriéndose con aceite de parafina, que reduce considerablemente la tasa de evaporación. De esta forma, se retrasó la cristalización y el crecimiento de 1/2 días a 3/4 días. Los cristales obtenidos mantenían la morfología habitual, aunque habían aumentado considerablemente el tamaño (hasta 0,8 mm de longitud) y carecían de defectos aparentes.
- **Cristalización a 37 °C.** Un aumento en la temperatura provoca mayor agitación y, generalmente, dificulta el crecimiento cristalino. Se realizaron ensayos de cristalización en gradientes de la condición original a 37 °C, con la intención de alterar el empaquetamiento, pero no se consiguieron cristales bajo ninguna condición.
- **Variación de volúmenes de gotas.** Es habitual modificar la relación entre la condición de cristalización y la gota de proteína, con la intención de explorar concentraciones que de otro modo no son evaluadas. Se observaron diferencias en la morfología de los cristales pero no se llegó a evitar la macia (Figura 5.13).



**Figura 5.13: Cristales de DacB<sub>61-238</sub> con variación de volúmenes.** Experimentos de cristalización de DacB<sub>61-238</sub> en 0,2 mM CaCl<sub>2</sub>, 0,1 M Tris-HCl pH 8,5 y 20 % (p/v) PEG 4000, variando los volúmenes de condición y proteína. Los paneles a – d se corresponden con las relaciones 1:2, 2:1, 1:3 y 3:1 (proteína:condición). A mayor concentración de proteína aumenta el tamaño de los cristales y los defectos. Todos los paneles están a la misma escala que el panel a.

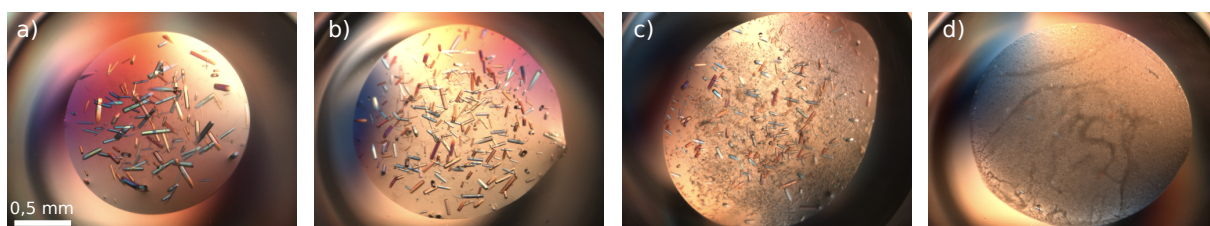


- **Utilización de diferentes sales.** El hecho de que se obtengan cristales en  $\text{CaCl}_2$  y  $\text{LiCl}$  puede implicar que otros cloruros también favorezcan la cristalización de DacB<sub>61–238</sub> pero alterando el empaquetamiento (MacRae & Doudna, 2007). Se ensayaron diferentes experimentos de cristalización en Tris-HCl pH 8,5 y concentraciones variables tanto de PEG 4000 como de  $\text{NH}_4\text{Cl}$ ,  $\text{BaCl}_2$ ,  $\text{CdCl}_2$ ,  $\text{CoCl}_2$ ,  $\text{CsCl}$ ,  $\text{CuCl}_2$ ,  $\text{FeCl}_2$ ,  $\text{MgCl}_2$ ,  $\text{MnCl}_2$ ,  $\text{NaCl}$ ,  $\text{NiCl}_2$ ,  $\text{KCl}$  y  $\text{ZnCl}_2$ , además de  $(\text{NH}_4)_2\text{SO}_4$ . Si bien algunas sales, como  $\text{NaCl}$  y  $\text{MgCl}_2$ , generaron numerosos cristales de pequeño tamaño, los intentos de escalado no fueron exitosos y devolvieron cristales con la morfología habitual.
- **Cristalización en presencia de detergentes y aditivos.** Se probaron paquetes comerciales de detergentes y aditivos (Tabla 3.1) con la intención de modificar la flexibilidad de zonas expuestas de DacB<sub>61–238</sub>, alterando las interacciones y el empaquetamiento cristalino. En general, o no se formaron cristales o crecieron con la morfología habitual. No obstante, es interesante el resultado obtenido al utilizar EDTA como aditivo, ya que se forma un precipitado cristalino filamentososo (Figura 5.14). Probablemente el EDTA secuestre el  $\text{Zn}^{2+}$  del sitio activo, desestabilizando el dominio catalítico y provocando su precipitación.



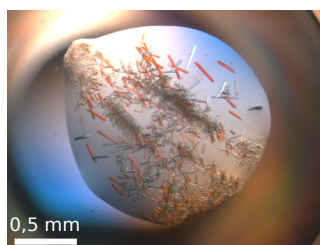
**Figura 5.14: Cristales de DacB<sub>61–238</sub> en presencia de EDTA.** El EDTA provoca la aparición de un precipitado cristalino, probablemente por el secuestro del  $\text{Zn}^{2+}$  del sitio activo y la desestabilización del dominio catalítico.

- **Utilización de MPD como aditivo.** El MPD (2-metil-2,4-pentanodiol) es un alcohol muy utilizado como aditivo en ensayos de cristalización ya que modifica la estabilidad de la proteína al rodearla y reducir su accesibilidad al solvente (Anand *et al.*, 2002). Se realizaron experimentos de cristalización en la condición original al 18 % (p/v) PEG 4000, añadiendo de un 5 a un 30 % (v/v) MPD. Los cristales resultantes mantenían la morfología habitual, decayendo a medida que aumentaba la concentración de MPD (Figura 5.15).



**Figura 5.15: Cristales de DacB<sub>61–238</sub> en presencia de MPD.** Experimentos de cristalización de DacB<sub>61–238</sub> en 0,2 mM  $\text{CaCl}_2$ , 0,1 M Tris-HCl pH 8,5 y 18 % (p/v) PEG 4000, en presencia de concentraciones crecientes de MPD. Los paneles a – d se corresponden con 5, 10, 15 y 20 % (v/v) MPD. Todos los paneles están a la misma escala que el panel a.

- **Cristalización en presencia de aminoácidos.** Un paquete comercial especialmente útil en ciertas ocasiones es el Morpheus (Tabla 3.1), que contiene mezclas desconocidas de aminoácidos, azúcares, alcoholes y otros compuestos de interés. Ya que DacB une fragmentos de origen peptídico, se ensayaron las condiciones con mezclas de aminoácidos, sin llegar a conseguir cristales.
- **Cristalización en presencia de fragmentos de peptidoglicano.** Se realizaron ensayos de cocrystalización e inmersión en fragmentos homogéneos de peptidoglicano correspondientes a los sustratos y productos de la proteína DacB, con alguna variación. En ningún caso se consiguió que el ligando apareciese unido a la estructura o modificase el empaquetamiento cristalino. En la Sección 5.4.6 se analizan con mayor detalle este tipo de ligandos.
- **Diferentes técnicas de sembrado.** Como se indicó en la Sección 3.1, el sembrado es una técnica habitual para mejorar el tamaño de los cristales y evitar defectos de crecimiento. Las técnicas de sembrado sobre DacB<sub>61–238</sub> se realizaron con la intención de separar núcleos de una única orientación, favoreciendo el crecimiento de un monocristal sin macla. En primer lugar, se realizó un sembrado de núcleos sobre un gradiente de PEG 4000 (14 % a 24 %), obteniendo cristales idénticos a los originales (Figura 5.16). Posteriormente se realizó un sembrado heterogéneo epitaxial, utilizando núcleos de DacB<sub>20–238</sub> sobre un gradiente de la condición de cristalización de DacB<sub>61–238</sub>. La finalidad era forzar la agregación de DacB<sub>61–238</sub> en una orientación distinta a la habitual, establecida previamente gracias a los cristales de DacB<sub>20–238</sub>, pero el resultado fue el habitual.



**Figura 5.16: Cristales de DacB<sub>61–238</sub> obtenidos por sembrado.** El sembrado de núcleos de DacB<sub>61–238</sub> devolvió cristales con la morfología habitual. En la gota se aprecia el surco dejado por el pelo de gato al liberar los núcleos, así como un gradiente de tamaño desde el surco hacia el exterior. Cerca del surco la liberación de núcleos ha sido mayor, por lo que se han formado muchos cristales de pequeño tamaño. En las zonas alejadas del surco hay menos núcleos, lo que favorece el crecimiento de cristales aislados y grandes.

Así pues, no fue posible alterar la morfología de los cristales de DacB<sub>61–238</sub> ni eliminar la macla. No obstante, estos experimentos permitieron obtener cristales de diferente calidad que, posteriormente, sirvieron para caracterizar la macla y mejorar progresivamente el modelo del dominio catalítico de DacB.

#### 5.3.4. Análisis de maclas y determinación estructural

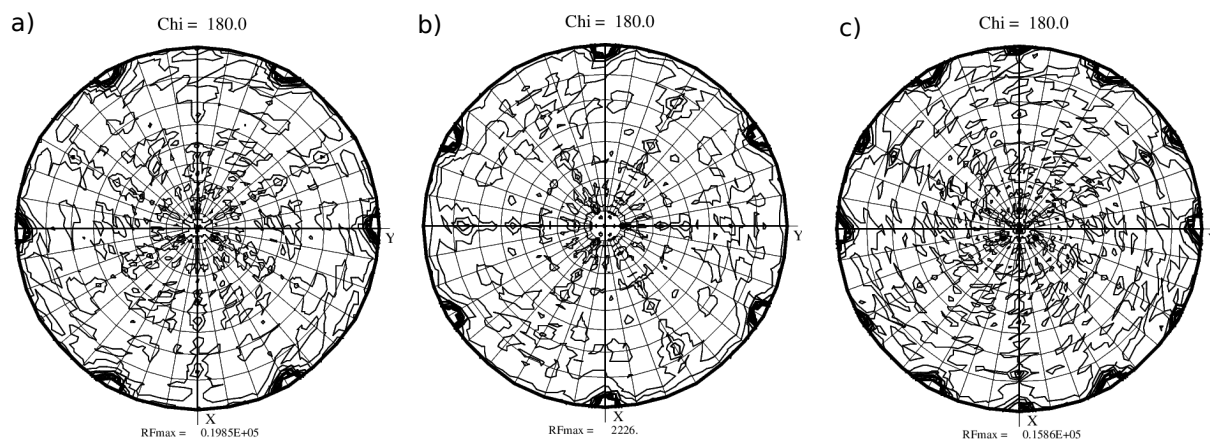
En la Tabla 5.4 se observó que existía cierta variabilidad en la geometría de cristales que, *a priori*, deberían ser iguales. En las siguientes líneas se describe hasta qué punto se ha podido acotar el problema tras los ensayos de optimización.

En primer lugar, se han observado dos tipos de celdilla unidad:

- **Celdilla corta:**  $a \simeq 37,8 \text{ \AA}$ ,  $b \simeq 37,8 \text{ \AA}$ ,  $c \simeq 111,2 \text{ \AA}$ ,  $\alpha = 90^\circ$ ,  $\beta = 90^\circ$ ,  $\gamma = 120^\circ$ .
- **Celdilla larga:**  $a \simeq 66,5 \text{ \AA}$ ,  $b \simeq 66,5 \text{ \AA}$ ,  $c \simeq 111,2 \text{ \AA}$ ,  $\alpha = 90^\circ$ ,  $\beta = 90^\circ$ ,  $\gamma = 120^\circ$ .

Como se indicó anteriormente, los reemplazos moleculares fallaban y las estadísticas no se correspondían con lo esperable. Los cristales optimizados presentaban grupos espaciales perteneciente al grupo puntual 32 (habitualmente  $P3_121$  y  $P3_112$ ). Los cristales de celdilla larga mostraban 1 molécula por unidad asimétrica (aunque sin ser un resultado sólido), mientras que los cristales de celdilla corta no tenían espacio para albergar ninguna molécula. Así pues, se decidió reducir la simetría y trabajar en el grupo puntual 3 ( $P3$ ,  $P3_1$  y  $P3_2$ ).

El análisis de las funciones de autorrotación de diferentes cristales llevó a la determinación de tres casos distintos, mostrados en la Figura 5.17. Los dos primeros casos (Figuras 5.17a y 5.17b) muestran una disposición de picos que se corresponde exactamente con un grupo de Laue  $\bar{3}m$  (312 y 321 respectivamente). En este punto se plantean dos hipótesis. En primer lugar, si el cristal realmente pertenece a un grupo puntual 3, los picos de las funciones de autorrotación indican que existe una macla hemihédrica que añade un eje binario al sistema. En cambio, si el cristal realmente no está maclado, nos encontraríamos frente a un grupo puntual 32 real. Ya que en el grupo puntual 32 no se resuelve la estructura y sí se encuentra solución en el grupo puntual 3, es de esperar que los cristales se comporten según la primera hipótesis (grupo puntual 3 con macla hemihédrica).



**Figura 5.17: Funciones de autorrotación de DacB<sub>61–238</sub>.** Calculadas con Molrep para  $\chi = 180^\circ$ .

**a)** Función de autorrotación correspondiente a un grupo de Laue  $\bar{3}m$ , de tipo 312. **b)** Función de autorrotación correspondiente a un grupo de Laue  $\bar{3}m$ , de tipo 321. **c)** Función de autorrotación correspondiente a un grupo de Laue  $6/mmm$ .

La tercera función de autorrotación se asemeja a la de un grupo de Laue  $6/mmm$  (Figura 5.17c). Así pues, es de esperar que el grupo espacial pertenezca al grupo puntual 622 o similar. La ausencia de solución de reemplazo molecular llevó a reducir la simetría del grupo puntual 622 al 32. Una solución en este grupo habría indicado la presencia de una macla hemihédrica, pero la falta de resultados concluyentes obligó a reducir aún más la simetría, llegando al grupo puntual 3, para el que sí se obtuvo una solución de reemplazo molecular aceptable. Por tanto, estos

cristales pertenecían al grupo puntual 3 y presentaban una macla tetartohédrica, siguiendo el comportamiento descrito en Yeates & Yu, 2008. Además, en estos casos se detectó una operación de pseudo-simetría traslacional, lo que coincide con el resto de casos descritos hasta la fecha (Roversi *et al.*, 2012). Así pues, la macla tetartohédrica presente en los cristales de DacB<sub>61–238</sub> es de tipo pseudo-merohédrico.

Mientras que en los primeros cristales de DacB<sub>61–238</sub> se observó mayoritariamente una macla tetartohédrica, a medida que se iban optimizando se consiguieron cristales con maclas hemihédricas. Esto implica que, si bien no se pudo eliminar por completo la macla tetartohédrica, se redujeron de cuatro a dos los dominios en los que se organizaban los cristales o, al menos, las fracciones de algunos dominios se vieron disminuidas hasta ser despreciables. En la Tabla 5.5 se muestran las diferentes combinaciones de leyes y fracciones de macla obtenidas en los mejores cristales. Si bien los casos de macla tetartohédrica son complejos debido a la división de las intensidades en cuatro contribuciones de distinto peso, los casos obtenidos con macla hemihédrica son perfectos o casi perfectos ( $\alpha \geq 0,5$ ), por lo que la reducción de cuatro dominios a dos dificultó aún más la resolución estructural.

**Tabla 5.5:** Fracciones de macla asociadas a cada ley de macla en DacB<sub>61–238</sub>.

MACLA	LEY DE MACLA		
	$-h, -k, l$	$h, -h - k, -l$	$-k, -h, -l$
Hemihédrica		0,45 – 0,50	
Hemihédrica			0,40 – 0,50
Tetartohédrica	0,10 – 0,20	0,30 – 0,45	0,10 – 0,20
Tetartohédrica	0,10 – 0,20	0,10 – 0,20	0,30 – 0,45

Además, fue bastante habitual encontrar cristales maclados en los que cada dominio difractaba a diferente resolución. Así pues, si bien los datos alcanzaban una resolución global de entre 1,6 y 1,8 Å, se realizaron varios procesamientos truncando los datos hasta resoluciones en torno a 2,5 Å, capa en la que las estimaciones de  $\alpha$  eran más precisas.

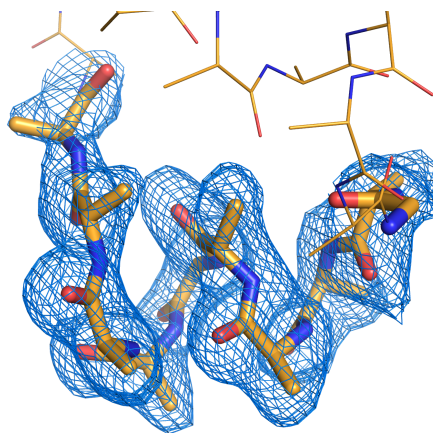
Al integrar los datos en el grupo espacial P3<sub>2</sub> (obtenido tras analizar las probabilidades y resultados en todos los posibles grupos espaciales pertenecientes al grupo puntual 3) se encontraron 1 y 3 moléculas por unidad asimétrica en las celdillas corta y larga respectivamente. Los reemplazos moleculares se realizaron con diferentes modelos, dependiendo de la calidad del cristal y del tipo de macla. Inicialmente se utilizó el modelo de DacB<sub>20–238</sub>, pero al ser un modelo incompleto e impreciso, se decidió continuar con los reemplazos moleculares utilizando la proteína VanXYg de *E. faecalis* (código PDB 4F78, Meziane-Cherif *et al.*, 2014).

Inicialmente sólo se disponía de cristales con macla tetartohédrica y celdilla corta, lo que suponía una molécula por unidad asimétrica. Durante el refinamiento de datos maclados se debe tener en cuenta tanto la ley de macla como su fracción  $\alpha$ , ya que es la manera de indicar a los algoritmos la relación existente entre los factores de estructura experimentales. En el momento de realizar la determinación estructural de DacB<sub>61–238</sub>, el programa phenix.refine (del conjunto

de programas PHENIX, Adams *et al.*, 2010) no contemplaba la posibilidad de incluir más de una ley de macla en el refinamiento. Únicamente la versión más reciente de Refmac5 (Murshudov *et al.*, 2011) aportaba esta opción. Así pues, se inició el refinamiento con Refmac5, incluyendo las tres leyes de macla y sus respectivas fracciones. No obstante, el modelo sufría repetidamente modificaciones en su geometría, sin poderse restringir de manera eficiente. Por tanto, se optó por una estrategia de refinamiento en phenix.refine, más conservador en el tratamiento de la geometría, aportando únicamente la ley de macla cuya fracción era mayoritaria (generalmente,  $-k, -h, -l$ ). Posteriormente, al obtener un modelo más completo y geométricamente optimizado, se volvería a Refmac5 para realizar ciclos de refinamiento incluyendo las tres leyes de macla.

Durante las primeras fases de construcción manual y refinamiento en phenix.refine, los factores  $R$  se situaron en torno a 0,40 – 0,50. Progresivamente se fue ajustando el modelo, llegando, finalmente, a valores de  $R_{\text{work}} \simeq 0,30$  /  $R_{\text{free}} \simeq 0,35$ . Para ello tuvieron que aplicarse restricciones de estructura secundaria y Ramachandran, además de aumentar ligeramente la contribución del modelo al proceso de refinamiento sobre la contribución de los factores de estructura experimentales. No obstante, el valor de  $R_{\text{free}}$  no se pudo estimar correctamente debido a que, durante la elección de las reflexiones aleatorias para el cálculo de este factor, no se tuvieron en cuenta dos de los cuatro dominios que componen el cristal. Tras conseguir un modelo que no era mejorable con phenix.refine, se realizaron ciclos de refinamiento en Refmac5, incluyendo las tres leyes de macla, aunque la mejora en el modelo y las estadísticas no fue muy sustancial. El modelo se fue mejorando progresivamente a medida que se iban optimizando los cristales. No obstante, en ningún caso se pudo obtener una estructura definitiva.

El motivo principal de la imposibilidad de obtener un modelo adecuado y completo es el sesgo que se produce en los mapas de datos maclados cuando el modelo no es bueno. Únicamente las zonas más intensas del mapa son fiables, mientras que el resto suelen presentar distinto grado de ambigüedad. Además, al utilizar una única ley de macla en un caso tetartohédrico, se pierde la información de dos dominios, por lo que los mapas resultantes aparecen aún más sesgados (Figura 5.18).

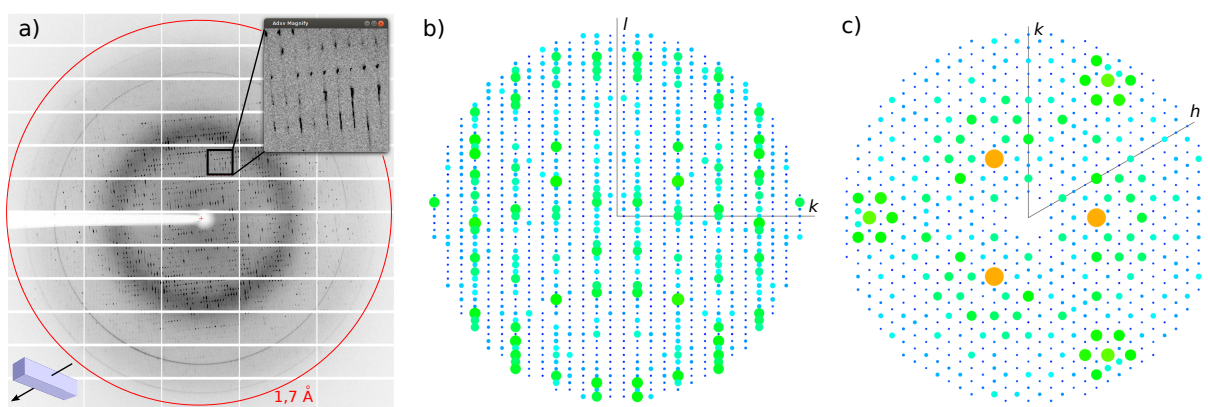


**Figura 5.18: Sesgo en los mapas de densidad electrónica de DacB<sub>61–238</sub>.** Mapa de densidad electrónica  $2F_{\text{obs}} - F_{\text{cal}}$  calculado a  $1,0 \sigma$ , tras 10 ciclos de refinamiento con un modelo de poli-Ala de DacB<sub>61–238</sub>. El mapa debería predecir las cadenas laterales de los residuos mostrados. En cambio, debido al sesgo, se ajusta perfectamente a las alaninas que se han modelado.

En el caso de  $\text{DacB}_{61-238}$ , la zona correspondiente al sitio de unión a zinc mantiene su intensidad aunque se modifique el modelo, mientras que los alrededores se van sesgando progresivamente. El sesgo en  $\text{DacB}_{61-238}$  surge por partir de un modelo inicial incompleto y con fallos, que no se ajusta bien a los factores de estructura experimentales, los cuales, debido a la macla, son débiles y contribuyen en menor medida al cálculo de la densidad electrónica. Así pues, la mayor contribución al cálculo de la densidad electrónica proviene de un modelo incompleto y con fallos, del que únicamente se puede asegurar la topología a nivel de estructura secundaria y las cadenas laterales de los residuos cercanos al sitio de unión a zinc. En general, el sesgo es difícil de detectar ya que los mapas suelen tener muy buena calidad en las zonas en las que se ha construido el modelo. De hecho, se debe sospechar de sesgo cuando un mapa presenta mejor calidad de la que debería en base a la calidad de los datos. Un buen mapa con factores  $R$  elevados puede indicar la presencia de sesgo. De igual manera, una zona está sesgada si al eliminar parte del modelo y refinar pocos ciclos se pierde la densidad en esa zona. En el caso de  $\text{DacB}_{61-238}$ , el porcentaje de mapa sesgado era bastante elevado, generando incertidumbre en la posición de determinadas cadenas laterales y en la geometría de lazos expuestos.

Por último, otro de los problemas era, como ya se indicó, la existencia de cristales con celdillas unidad de distintos tamaños. En general, los patrones de difracción de todos los cristales, optimizados o no, presentan numerosos daños. Es un rasgo común a todos ellos la aparición de líneas ordenadas entre reflexiones intensas, por lo que la indexación podría fallar en determinados casos (Figura 5.19a). Además, analizando cortes bidimensionales del espacio recíproco de cristales con celdilla larga se puede observar un patrón de reflexiones intensas intercaladas entre reflexiones más débiles (Figuras 5.19b y 5.19c), hecho que puede deberse a problemas durante el indexado.

Los intentos de forzar el integrado en una celdilla distinta a la más probable para cada cristal han fallado, devolviendo estadísticas desfavorables en cada caso. Así pues, la existencia de una

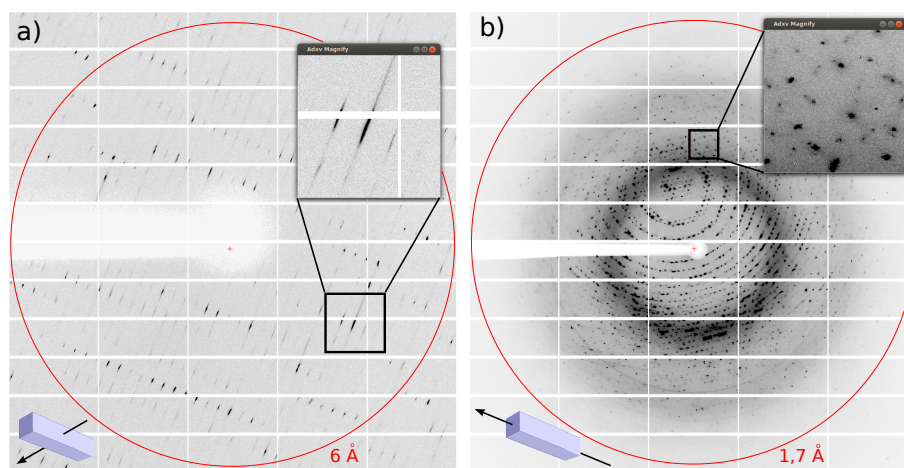


**Figura 5.19: Patrón de difracción de  $\text{DacB}_{61-238}$  y espacio recíproco. a)** Patrón de difracción habitual de los cristales de  $\text{DacB}_{61-238}$ , mostrando una serie de reflexiones puntuales y líneas continuas. En la esquina inferior izquierda se muestra un diagrama de la orientación del haz de rayos X respecto al cristal. **b)** Corte del espacio recíproco de un cristal de celdilla larga, mostrando el plano  $(l, k)$ . Se observa un aumento de intensidad en las reflexiones cada tres capas a lo largo del eje  $k$ . **c)** Corte del espacio recíproco de un cristal de celdilla larga, mostrando el plano  $(h, k)$ . Se observa un aumento de intensidad en las reflexiones cada tres capas a lo largo de ambos ejes.



celdilla larga y una corta puede deberse bien a la existencia de dos tipos distintos de cristal, bien a los defectos acumulados durante el crecimiento de un sólo tipo de cristal, alterando el indexado hacia una celdilla corta o larga dependiendo de cada caso particular.

No obstante, se planteó una hipótesis adicional, en la que las líneas del patrón de difracción no fueran líneas sino el reflejo de la existencia de un eje cristalográfico muy largo en la celdilla. Un eje de varios cientos de ångströms da lugar a reflexiones muy juntas en el patrón. Si durante el experimento de difracción se recogen datos a alta resolución, es probable que las reflexiones cercanas se solapen, dando lugar a líneas. Ya que los cristales de DacB<sub>61–238</sub> se habían recogido siempre a una resolución aproximada de 1,5 Å, se realizó una toma de datos a 6 Å, con la intención de separar las reflexiones a baja resolución y analizar la posible existencia de solapamiento. Como se puede observar en la Figura 5.20a, las líneas se mantienen pese a haber alejado el detector del cristal, por lo que se descarta la hipótesis del eje cristalográfico largo.



**Figura 5.20: Defectos en los patrones de difracción de DacB<sub>61–238</sub>.** a) Patrón de difracción a 6 Å, recogido con la orientación habitual (esquina inferior izquierda), mostrando que las líneas se mantienen constantes y descartando la existencia de un eje cristalográfico largo. b) Patrón de difracción recogido con una orientación del haz de rayos X distinta a la habitual (esquina inferior izquierda). La ampliación muestra el desdoblamiento de reflexiones debido a la acumulación de daños a lo largo del eje largo del cristal.

Además, ya que la orientación del cristal respecto al haz de rayos X era constante en todos los experimentos de difracción, se decidió recoger algún conjunto de datos orientando el cristal de forma que el haz de rayos X fuera paralelo al eje largo del cristal. Como era de esperar, el patrón recogió la acumulación de defectos producidos a lo largo de dicha dirección, mucho mayor de la que se produce si el haz atraviesa un fragmento más corto del cristal (Figura 5.20b). Por tanto, se descartó el uso de esta estrategia durante los siguientes experimentos de difracción.

En la Tabla 5.6 se muestran las estadísticas del mejor modelo de DacB<sub>61–238</sub>, obtenido a partir de un conjunto de datos de celdilla corta con macla pseudo-merohédrica tetartohédrica con fracciones de macla aproximadas de 0,10, 0,10 y 0,40, siendo la mayoritaria la correspondiente a la ley  $h, -h-k, -l$ . Como se puede observar, tanto los factores  $R$  como los valores de Ramachandran han mejorado significativamente con respecto al modelo de DacB<sub>20–238</sub> (Tabla 5.3). No obstante, como se ha indicado a lo largo de esta sección, el modelo sigue siendo parcial, con zonas dudosas y probablemente sesgadas.

**Tabla 5.6:** Estadísticas de refinamiento de DacB<sub>61–238</sub>.

PARÁMETRO	DACB <sub>20–238</sub>
Resolución (Å)	32,1 – 2,2
$R_{\text{work}}/R_{\text{free}}$	0,31 / 0,35
Nº de átomos	
Proteína (no H)	1028
Zinc	1
Agua	0
Factores B (Å <sup>2</sup> )	
Proteína	22,3
Zinc	17,9
RMSD de enlaces	
Distancias (Å)	0,010
Ángulos (°)	2,068
Ramachandran (%)	
Favorables	72,7
Desfavorables	5,5

### 5.3.5. Faseado experimental de cristales maclados

Como se comentó en la Sección 3.3.4, resulta más sencillo resolver datos maclados mediante reemplazo molecular que mediante faseado experimental. No obstante, existen trabajos en los que se han resuelto, mediante SAD o MAD, cristales con maclas hemihédricas con fracciones de macla cercanas a 0,5 o incluso con maclas pseudo-merohédricas tetartohédricas (Dauter, 2003; Rudolph *et al.*, 2003; Barends *et al.*, 2005). Si bien son casos excepcionales de muy elevada dificultad, no conviene descartar este tipo de experimentos para mejorar las fases de cristales maclados, más aún cuando se dispone de un modelo de reemplazo molecular que se podría utilizar en un procesado por MR-SAD.

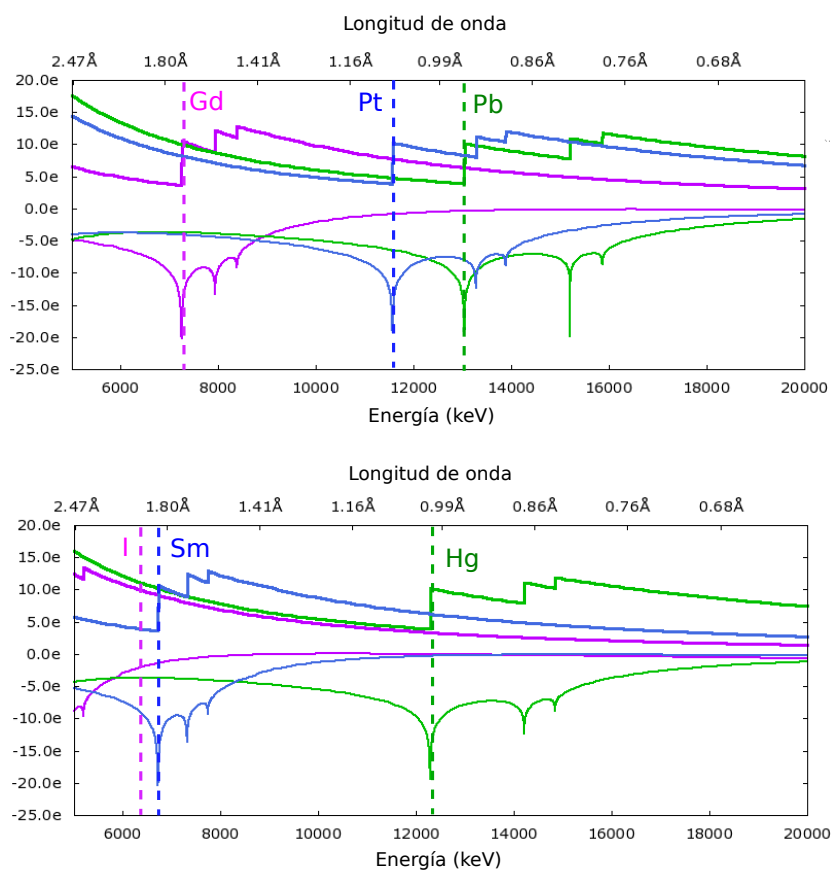
Si bien la estrategia más frecuente es intentar un experimento de Se-SAD mediante la inclusión de selenometioninas o selenocisteínas en la proteína, en el caso de DacB<sub>61–238</sub> se descartó por la presencia de únicamente 2 residuos de metionina y ninguna cisteína en 178 aminoácidos. En condiciones normales, ante la falta de otra estrategia, se puede intentar una aproximación bajo esta situación, pero gracias al modelo parcial de DacB se observó que una de las metioninas (Met105) podía estar expuesta al solvente, lo que puede disminuir notablemente su contribución a la dispersión anómala.

Por tanto, se realizaron experimentos de inmersión (a diferentes concentraciones y tiempos) en soluciones con diferentes dispersores anómalos, esperando que alguno penetrara en el cristal y permitiera la obtención de fases mediante SAD o MAD (Sección 3.3.3.4). La elección de los dispersores anómalos se realizó, en parte, siguiendo el criterio del servidor HATODAS (*Heavy-atom Database System*; Sugahara *et al.*, 2005), que estima, en base a secuencia y a comparación con la literatura, los dispersores anómalos que se pueden unir a la proteína problema. De este modo se consideró la utilización de gadolinio, mercurio, samario, platino y plomo. Además, se decidió utilizar también yodo, un elemento muy frecuente en los experimentos de SAD (Dauter & Dauter, 2007; Abendroth *et al.*, 2011). Por último, en relación con el yodo, se utilizó el



denominado *triángulo mágico* (ácido 5-amino-2,4,6-triiodoisoftálico, I3C; Beck *et al.*, 2008), sustituyente de las aguas del solvente y altamente afín tanto por los grupos carbonilo de la cadena principal como por las cadenas laterales de residuos tipo Asp, Glu, Lys, Ser, Tyr o Thr.

Los dispersores elegidos abarcan un rango amplio de energías (Figura 5.21). Salvo el yodo, cuya banda  $L_I$  de absorción se sitúa cerca de los 5 keV, el resto de elementos no baja de 6 keV. Para cada dispersor anómalo, exceptuando al yodo, se recogieron datos de difracción en el pico de absorción de su banda  $L_{III}$ , en la correspondiente inflexión y en una región remota, con la intención de realizar tanto un SAD como un MAD. En el caso del yodo, únicamente se puede recoger un conjunto de datos a la menor energía posible (siendo el factor limitante el rango de energías de la línea de radiación sincrotrón), por lo que la estrategia a seguir con este elemento debe ser un SAD. Además, en todos los experimentos de difracción se recogieron un total de  $360^\circ$ , lo que eleva la multiplicidad y puede mejorar la determinación de las diferencias anómalas.



**Figura 5.21:** Espectros teóricos de absorción de los dispersores anómalos usados en el faseado de DacB<sub>61–238</sub>. En líneas punteadas se muestran las longitudes de onda a las que se han tomado los conjuntos de datos correspondientes al pico de absorción. En todos los casos, salvo en el yodo, se corresponden con la banda  $L_{III}$  de absorción.

En la Tabla 5.7 se muestran las estadísticas más significativas, relacionadas con el análisis de la señal anómala, de los conjuntos de datos tomados en el pico de absorción del I3C. Como se puede observar, únicamente los cristales inmersos en la solución de I3C generaron suficiente señal anómala como para avanzar en la determinación estructural por SAD. No obstante, la señal anómala no fue lo suficientemente fuerte como para generar un contraste que permitiese distinguir la subestructura original de la inversa o trazar un modelo inicial fiable. Los conjuntos

de datos de cristales inmersos en el resto de elementos no llegaron a mostrar un  $\langle d''/\text{sig} \rangle$  superior a 1,3 (umbral mínimo para considerar la existencia de señal anómala significativa) en ninguna capa de resolución para el pico de absorción, impidiendo la resolución por SAD. Los intentos de aplicar la metodología de MAD tampoco dieron resultados positivos.

**Tabla 5.7:** Estadísticas relacionadas con el faseado experimental de DacB<sub>49–238</sub>. Todos los cristales alcanzaron una resolución aproximada de 1,8 Å, fueron integrados en el grupo espacial P3<sub>2</sub> y presentaban macla merohédrica hemihédrica con fracciones de macla superiores a 0,40.

PARÁMETRO	DacB <sub>61–238</sub> : I3C
Longitud de onda (Å)	1,8960
Resolución (Å)*	3,5
$CC_{\text{anom}}$	0,42
$CC_{\text{all}}/CC_{\text{weak}}$	0,32 / 0,10
CFOM	0,42
Sitios**	5
Contraste directa (%)	59,9
Contraste inversa (%)	61,2
$CC_{\text{map}}$ directa ***	0,71
$CC_{\text{map}}$ inversa ***	0,73

\* El límite de resolución de la señal anómala se establece para  $\langle d''/\text{sig} \rangle = 1,3$ .

\*\* Se consideran los que presentan una ocupación superior al 70 %.

\*\*\* Calculado para la máxima resolución del conjunto de datos.

Así pues, el modelo obtenido por reemplazo molecular, cuyas estadísticas se muestran en la Tabla 5.6, ha sido el mejor obtenido tras los intentos de eliminar las maclas y mejorar las fases descritos anteriormente.

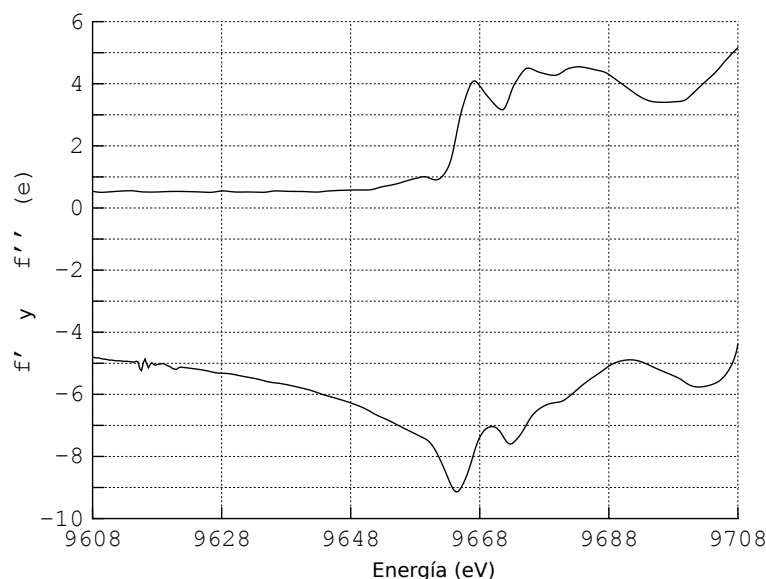
#### 5.3.6. Determinación de la existencia del sitio de unión a zinc

Como se ha comentado con anterioridad, se sospechaba de la existencia de un sitio de unión a zinc debido tanto al análisis del plegamiento de DacB<sub>20–238</sub> como a las predicciones basadas en secuencia realizadas sobre la proteína VanXYg de *E. faecalis*.

Dado que en ningún punto de la purificación y cristalización de DacB<sub>61–238</sub> se añadieron compuestos con zinc, es de esperar que un espectro de fluorescencia en torno al máximo de absorción de este elemento permita determinar si la proteína contiene zinc. Como se puede observar en la Figura 5.22, el espectro muestra un pico coincidente con el máximo de absorción del zinc, lo que indica la presencia de este elemento en el cristal. Así pues, se confirma que la proteína DacB incluye zinc de forma natural.

#### 5.3.7. Análisis estructural

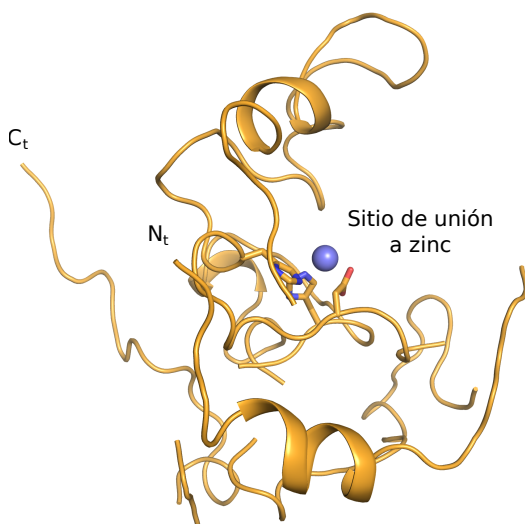
El modelo parcial de DacB<sub>61–238</sub> abarca desde el residuo Glu90 hasta el Glu232, aunque incluye zonas de poli-Ala y cortes en la secuencia de determinadas regiones. Además, la existencia de sesgo en los mapas impide tener una confianza plena en las coordenadas atómicas de parte del modelo.



**Figura 5.22:** Espectro de fluorescencia de DacB<sub>61–238</sub> en torno al zinc. La presencia de un pico a la energía de absorción del zinc indica la presencia de este elemento en el cristal.

#### 5.3.7.1. Estructura general

La estructura de DacB<sub>61–238</sub> muestra el plegamiento del dominio catalítico de DacB con una topología coincidente con la observada en los datos incompletos y a baja resolución de DacB<sub>20–238</sub> (Sección 5.2.4.1). Nuevamente se observa el sitio de unión a zinc sobre la lámina  $\beta$ , rodeado por dos grupos de hélices  $\alpha$  (Figura 5.23), formado por los residuos His153, Asp160 e His207.



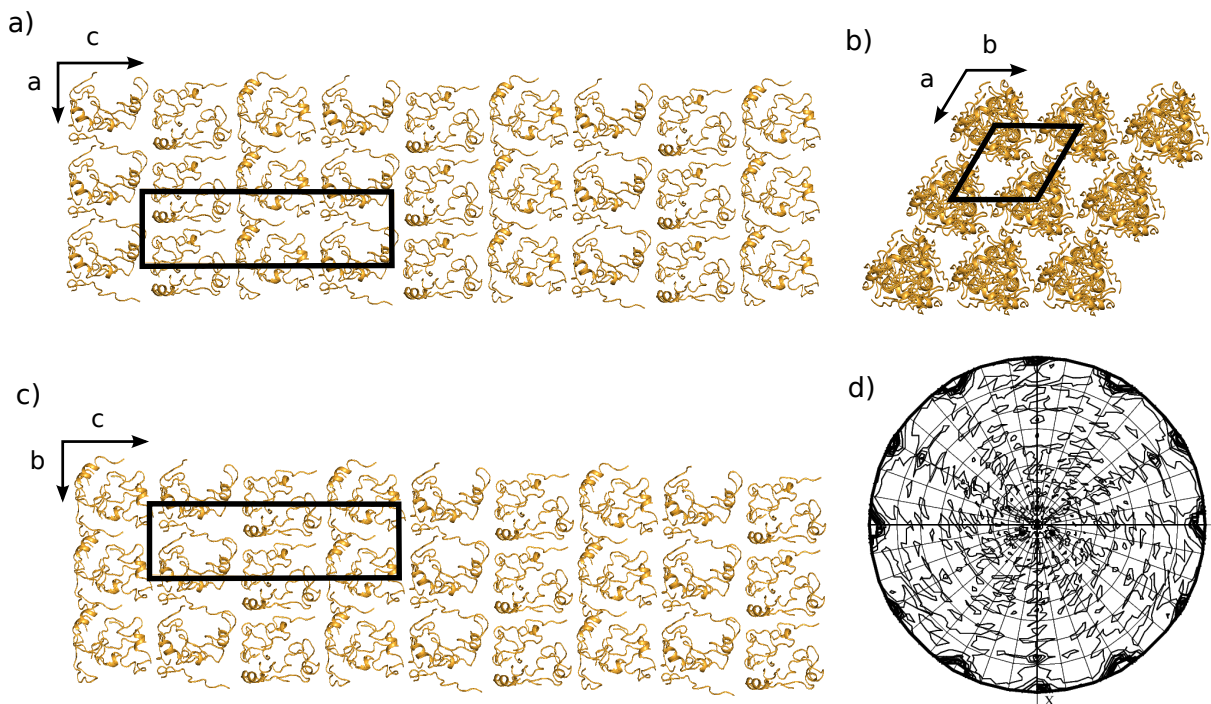
**Figura 5.23:** Estructura general de DacB<sub>61–238</sub>. Se mantiene el plegamiento observado en DacB<sub>20–238</sub>, consistente en una lámina  $\beta$  rodeada de hélices  $\alpha$ . Existen numerosos cortes en la secuencia debido a la imposibilidad de determinar con eficiencia las zonas más flexibles del modelo. El átomo de  $\text{Zn}^{2+}$  se representa en esferas.

Si bien este modelo presenta mejores factores  $R$  de refinamiento que el modelo de DacB<sub>20–238</sub>, el sesgo debido a la presencia de maclas evita que se puedan determinar con precisión las posiciones de las cadenas laterales de gran parte del modelo. Este hecho, junto con la ausencia

de densidad en varios lazos que conectan elementos de estructura secundaria, puede ser motivo de desfases en la posición de los aminoácidos que componen algunas zonas de la estructura, llevando a factores  $R$  elevados. El otro motivo por el que los factores  $R$  pueden ser altos es porque los datos de difracción de un patrón maclado no son lo suficientemente precisos como para obtener un modelo más fiable que el actual. No obstante, al desconocer la estructura real de la proteína, no se puede determinar si la falta de precisión del modelo de DacB<sub>61–238</sub> se debe a uno de los dos casos o a la combinación de ambos.

### 5.3.7.2. Empaquetamiento cristalino

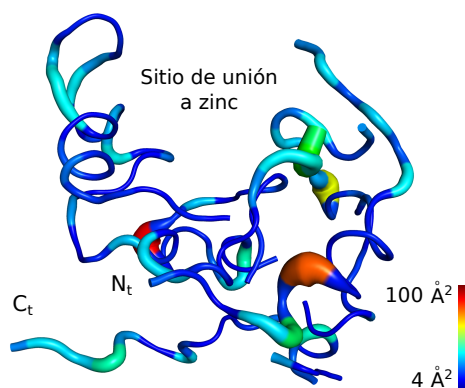
La organización de las moléculas de DacB<sub>61–238</sub> en el cristal se corresponde con la esperable para un grupo puntual 3, siendo uno de los ángulos 120°. Los planos  $(a, c)$  y  $(b, c)$  (Figuras 5.24a y 5.24b respectivamente) muestran un empaquetamiento muy compacto, lo que explicaría la alta resolución que alcanzan estos cristales. La macla tetartohédrica se refleja en la función de autorrotación (Figura 5.24d), que se corresponde con un grupo de Laue  $6/mmm$  en lugar de con un grupo  $\bar{3}$ , como se explicó en la Sección 5.3.4.



**Figura 5.24: Simetría y empaquetamiento de DacB<sub>61–238</sub>.** El recuadro negro hace referencia a la celdilla unidad, indicándose además la orientación de cada eje en cada caso. **a)** Plano  $(a, c)$ . **b)** Plano  $(a, b)$ . **c)** Plano  $(b, c)$ . **d)** Función de autorrotación calculada para  $\chi = 180^\circ$  mediante Molrep, mostrando los picos característicos de un grupo puntual  $6/mmm$ .

### 5.3.7.3. Agitación térmica

Otra característica asociada al estrecho empaquetamiento y a la alta resolución es la baja agitación térmica. El modelo de DacB<sub>61–238</sub> presentó valores B promedio de entre 4 y 40 Å<sup>2</sup>, si bien algunos residuos presentaron una agitación inusualmente elevada debido a una falta de densidad precisa que justificase su correcta ubicación (Figura 5.25).



**Figura 5.25: Representación de factores de agitación térmica de DacB<sub>61-238</sub>.** Los valores cercanos al azul reflejan baja agitación térmica, mientras que un virado hacia el rojo implica mayor agitación térmica y mayor flexibilidad. Por motivos de claridad, se ha indicado la posición del sitio de unión a zinc.

#### 5.3.7.4. Otras moléculas y solvente

Además del  $\text{Zn}^{2+}$  presente en el sitio de unión a zinc, se intentaron modelar moléculas de agua alrededor de la estructura. No obstante, pese a la alta resolución de los mapas, el sesgo producido por la macla impidió que se pudieran trazar elementos del solvente.

### 5.4. Tercera construcción (DacB<sub>49-238</sub>)

El crecimiento tan rápido de los cristales de DacB<sub>61-238</sub> (24/48 horas), su sólida apariencia física y la alta resolución alcanzada en los experimentos de difracción (hasta 1,6 Å) indican que la construcción tiende a empaquetarse con facilidad, generando cristales de alta calidad que llevarán a una estructura estable y detallada. Además, la falta de éxito en los numerosos intentos de eliminar la macla durante el proceso de cristalización sugiere que es necesaria una aproximación más radical.

Debido a ello, se decidió realizar una nueva construcción, añadiendo algunos aminoácidos al extremo N-terminal de DacB<sub>61-238</sub>. La finalidad de esta aproximación era desestabilizar ligeramente el empaquetamiento de la construcción, entorpeciendo la dinámica de agregación y desfavoreciendo la formación de diferentes orientaciones en el cristal. Así pues, se diseñó una nueva construcción, DacB<sub>49-238</sub> (de Gly49 a Asp238), que contenía 12 residuos más que DacB<sub>61-238</sub>, pertenecientes a la región predicha como flexible (Figura 5.26). En total, 190 aminoácidos, con un peso molecular teórico de 21,3 kDa y un pI teórico de 5,0.

#### 5.4.1. Expresión y purificación

La obtención de DacB<sub>49-238</sub> fue realizada por el grupo del Prof. Sven Hammerschmidt, del Departamento de Genética Microbiana de la Universidad de Greifswald, Alemania, siguiendo el protocolo descrito para DacB<sub>20-238</sub> (Sección 5.2.1). La proteína final se obtuvo a una concentración de 15 mg/mL en 20 mM Tris-HCl pH 8,5.

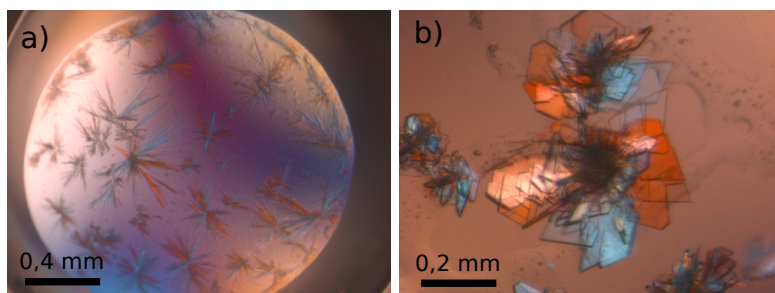
10	20	30	40	50	60
MKKRYLVLT	LLALSLAACS	QEKTKNEDGE	TKTEQTAKAD	GTVGSKSQGA	AQKKA <del>EVVNK</del>
70	80	90	100	110	120
GDYYSIQGKY	DEIIVANKHY	PLSKDYNPGE	NPTAKAELVK	LIKAMQEAGF	PISDHYSGFR
130	140	150	160	170	180
SYETQTKLYQ	DYVNQDGKAA	ADRYSARPGY	SEHOTGLAFD	VIGTDGDLVT	EEKAAQWLLD
190	200	210	220	230	
HAADYGFVVR	YLGKKEKETG	YMAEEWHLRY	VGKEAKEIAA	SGLSLEEYYG	FE <del>GGDYVD</del>

**Figura 5.26: Secuencia de la construcción DacB<sub>49–238</sub>.** En gris se muestra el péptido señal, mientras que en rojo se muestra el dominio catalítico. El sombreado amarillo se corresponde con la secuencia de la construcción de DacB<sub>49–238</sub>.

#### 5.4.2. Cristalización

La cristalización de DacB<sub>49–238</sub> se inició con los ensayos habituales de cristalización mediante técnicas de alto rendimiento, como se describieron en la Sección 5.2.2. Además, se realizaron ensayos de cristalización en variaciones de las condiciones de cristalización tanto de DacB<sub>61–238</sub> (CaCl<sub>2</sub>, Tris-HCl pH 8,5 y PEG 4000) como de DacB<sub>20–238</sub> (ZnCl<sub>2</sub>, Tris-HCl pH 8,0 y PEG 6000, en presencia y ausencia de CYMAL-3).

Sorprendentemente, DacB<sub>49–238</sub> cristalizó en 10 mM ZnCl<sub>2</sub>, 0,1 M Tris-HCl pH 8,0 y 14 % (p/v) PEG 6000, al igual que lo hizo DacB<sub>20–238</sub>, dando lugar al mismo hábito cristalino pero con mayor solidez y calidad óptica (Figura 5.27). En esta ocasión, no se observaron cristales en presencia de CYMAL-3, por lo que su papel en la cristalización de DacB<sub>20–238</sub> podría estar únicamente asociado a la estabilización de la región flexible ausente en DacB<sub>49–238</sub> (Ser20 – Gly49).



**Figura 5.27: Cristales de DacB<sub>49–238</sub>.** a) Cristales sin optimizar, en una gota de 2  $\mu$ L. b) Cristales optimizados, en una gota de 2  $\mu$ L.

#### 5.4.3. Determinación estructural

Los mejores cristales obtenidos en la etapa anterior se crioprotegieron en 20 % (v/v) glicerol y se llevaron a la línea I03 del DLS para ser difractados. Las estadísticas de difracción se recogen en la Tabla 5.8. El indexado, integrado y escalado se realizó con XDS y Aimless.

La celdilla unidad contenía 2 moléculas por unidad asimétrica, con un contenido de solvente del 46,4 % y un coeficiente de Matthews de 2,29 Å<sup>3</sup>/Da. Para el reemplazo molecular, realizado

**Tabla 5.8:** Datos y estadísticas de difracción de DacB<sub>49–238</sub>.

PARÁMETRO	DACB <sub>49–238</sub>
Grupo espacial	P2 <sub>1</sub>
$a, b, c$ (Å)	51,95, 37, 63, 102, 84
$\alpha = \beta = \gamma$ (°)	90,0, 97,98, 90,0
Resolución (Å)	101,8 – 2,0 (2,05 – 2,00)
Longitud de onda (Å)	1,0000
Línea	I03 (DLS)
Temperatura (K)	100
Nº de reflexiones	
Totales	178236
Únicas	26483
$\langle I/\sigma I \rangle$	8,2 (2,3)
$R_{\text{pim}}$	0,08 (0,33)
$CC_{1/2}$	1,00 (0,79)
Complejidad (%)	99,7 (99,2)
Multiplicidad	6,7 (6,6)

Los valores entre paréntesis hacen referencia a la última capa de resolución.

con Phaser, se utilizó como modelo la estructura parcial construida a partir de los datos maclados de DacB<sub>61–238</sub>. El resultado mostraba discrepancias entre la densidad electrónica calculada tras el primer refinamiento y numerosas cadenas laterales de ciertas zonas del modelo, aunque la cadena principal parecía, en general, acomodarse bastante bien a la densidad. Ante la posibilidad de un desfase de algunos residuos en la cadena de DacB<sub>61–238</sub> respecto a la cadena de DacB<sub>49–238</sub> (debido en parte a las condiciones en las que se construyó el modelo de DacB<sub>61–238</sub>) y gracias a la buena calidad de las primeras fases calculadas para la cadena principal tras el reemplazo molecular, se decidió priorizar la construcción automática de DacB<sub>49–238</sub> en Buccaneer, evitando la introducción de sesgo durante los ciclos de refinamiento iniciales. Para ello, se tomó el mapa de densidad electrónica calculado por Phaser y un modelo de DacB<sub>61–238</sub> truncado, manteniendo únicamente el sitio de unión a zinc (ya que su ubicación, composición y orientación eran fiables). Tras realizar algunos ciclos de extensión del modelo y mejora de las fases, el mapa resultante se introdujo en ARP/wARP, más eficiente a alta resolución, para realizar una construcción de la estructura sin modelo previo, teniendo como única información las fases calculadas y la secuencia de DacB<sub>49–238</sub>. De esta forma, se obtuvo aproximadamente el 90 % de DacB<sub>49–238</sub> perfectamente construido en un tiempo significativamente menor del que hubiera supuesto reconstruir a mano la estructura a partir de la solución de reemplazo con DacB<sub>61–238</sub>.

Finalmente, se realizaron varios ciclos de refinamiento en Refmac5, incluyendo refinamiento de NCS (un único grupo, correspondiente a cada monómero de DacB<sub>49–238</sub>) y restricciones de distancias de coordinación, seguidos de revisión y reconstrucción manual en Coot. Las estadísticas de refinamiento se muestran en la Tabla 5.9.

Tanto los factores B de agitación térmica como los valores de RMSD de enlaces y de Ramachandran son los esperables para unos datos de la calidad de los obtenidos con DacB<sub>49–238</sub>, lo que indica buena correlación entre la estructura y los datos de difracción.

**Tabla 5.9:** Estadísticas de refinamiento de DacB<sub>49–238</sub>.

PARÁMETRO	DACB <sub>49–238</sub>
Resolución (Å)	101,84 – 2,0
$R_{\text{work}}/R_{\text{free}}$	0,18 / 0,24
Nº de átomos	
Proteína (no H)	2950
Zinc	13
Fosfato	10
Agua	479
Factores B (Å <sup>2</sup> )	
Proteína	23,7
Zinc	26,8
Fosfato	33,7
RMSD de enlaces	
Distancias (Å)	0,015
Ángulos (°)	1,802
Ramachandran (%)	
Favorables	97,3
Desfavorables	0,3
<b>Código PDB</b>	<b>4D0Y</b>

#### 5.4.4. Análisis estructural

La calidad de los mapas de densidad electrónica permitió construir las dos moléculas de la unidad asimétrica de DacB<sub>49–238</sub> desde el residuo Ala55 hasta el extremo C-terminal, Asp238, lo que finalmente supone un modelo completo y detallado del dominio catalítico de DacB.

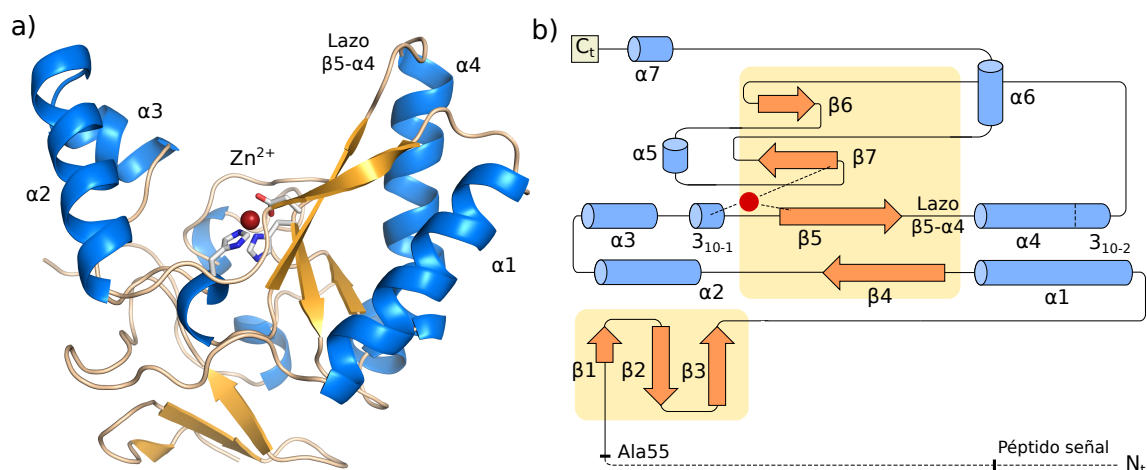
##### 5.4.4.1. Estructura general

En la región N-terminal de DacB<sub>49–238</sub> se sitúa una pequeña lámina  $\beta$ , formada por 3 hebras  $\beta$  antiparalelas ( $\beta_1 - \beta_3$ ), que conecta el dominio catalítico propiamente dicho con la región flexible característica de DacB (Ser20 – Ala55). El dominio catalítico de DacB muestra un plegamiento globular de tipo  $\alpha + \beta$ , con un núcleo de 4 hebras  $\beta$  antiparalelas ( $\beta_4 - \beta_8$ ) rodeadas de 7 hélices  $\alpha$  ( $\alpha_1 - \alpha_7$ ) y 2 hélices  $3_{10}$  ( $3_{10-1}$  y  $3_{10-2}$ ) (Figura 5.28). Este plegamiento se corresponde con el observado en DacB<sub>20–238</sub> y DacB<sub>61–238</sub>, consistente en dos grupos de hélices entre las cuales se encuentra el sitio de unión a zinc, sobre una lámina  $\beta$ . Además, como se verá más adelante, existe un lazo de especial interés situado entre hebra  $\beta_5$  y la hélice  $\alpha_4$ .

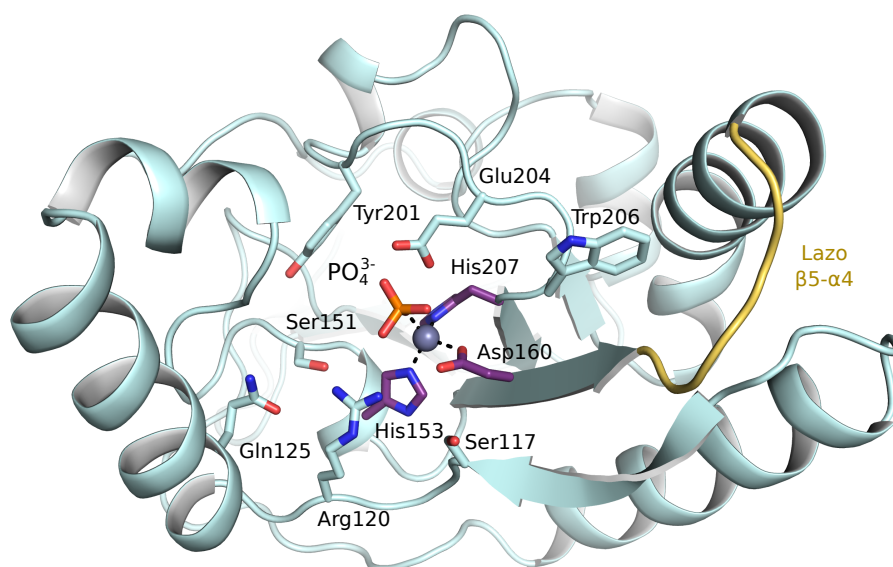
Como se observó en los modelos parciales de DacB<sub>20–238</sub> y DacB<sub>61–238</sub>, el sitio catalítico consta de un sitio de unión a zinc, formado por los residuos His153 ( $3_{10-1}$ ), Asp160 ( $\beta_5$ ) e His207 ( $\beta_7$ ). La densidad electrónica ha permitido localizar con precisión la posición del  $\text{Zn}^{2+}$  en el sitio y de un fosfato que completa la coordinación, probablemente arrastrado desde el proceso de purificación (Figura 5.29).

Las dos cadenas muestran el mismo plegamiento y una alta similitud (RMSD = 0,166 Å), aunque existen diferencias en la orientación y composición del lazo que conecta la hebra  $\beta_5$  y la hélice  $\alpha_4$  (lazo  $\beta_5 - \alpha_4$ ). Mientras que en una cadena el lazo es más corto (Thr164 – Asp167)





**Figura 5.28: Estructura general de DacB<sub>49-238</sub>.** **a)** Estructura general de DacB<sub>49-238</sub> mostrando los elementos de estructura secundaria (etiquetando las hélices principales) y el sitio de unión a zinc (varillas blancas y esfera roja). **b)** Diagrama de topología de DacB<sub>49-238</sub>, indicando la ubicación del sitio de unión a zinc (círculo rojo y líneas punteadas).

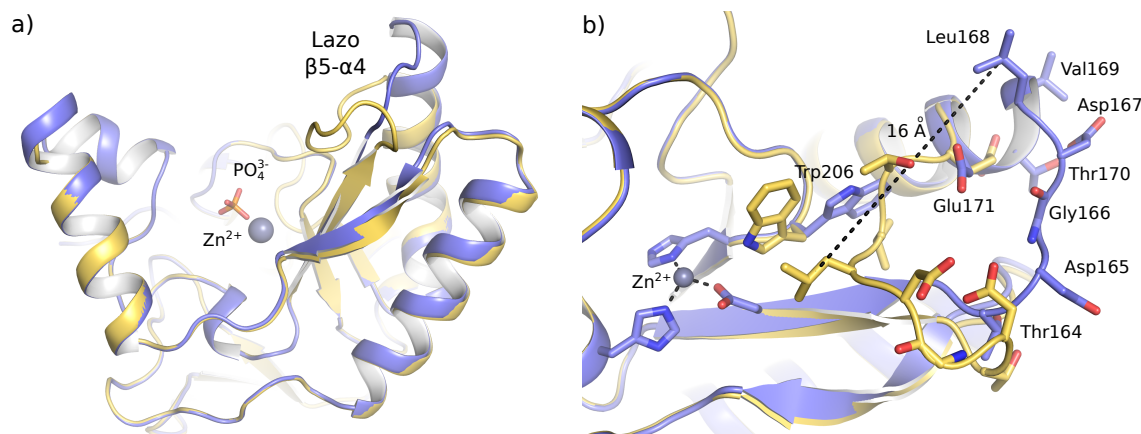


**Figura 5.29: Sitio de unión a zinc de DacB<sub>49-238</sub>.** Se muestra el átomo de  $\text{Zn}^{2+}$  (esfera) coordinado con la His153, el Asp160, la His207 (varillas moradas) y el fosfato ( $\text{PO}_4^{3-}$ , varillas naranjas). Además, se han representado en varillas azules otros residuos importantes en la catálisis por su cercanía al sitio de unión a zinc. El lazo  $\beta_5 - \alpha_4$  se representa en amarillo.

y se encuentra orientado hacia el solvente, en la otra cadena parte de la hélice  $\alpha_4$  pierde su estructura y añade 4 residuos al lazo, de mayor tamaño (Thr164 – Glu171) y orientado hacia el sitio catalítico. Como resultado de esta reorientación, el sitio catalítico sufre una modificación significativa de su tamaño, principalmente por el desplazamiento del Trp206. Así pues, se puede hablar de una conformación abierta y una conformación cerrada, en función de si el lazo  $\beta_5 - \alpha_4$  se encuentra orientado hacia el exterior o hacia el sitio catalítico respectivamente (Figura 5.30).

#### 5.4.4.2. Empaquetamiento cristalino

La construcción DacB<sub>49-238</sub> cristalizó en el grupo espacial monoclinico  $\text{P2}_1$ , incluyendo dos moléculas por unidad asimétrica. Así pues, hay un total de 4 moléculas por celdilla unidad.



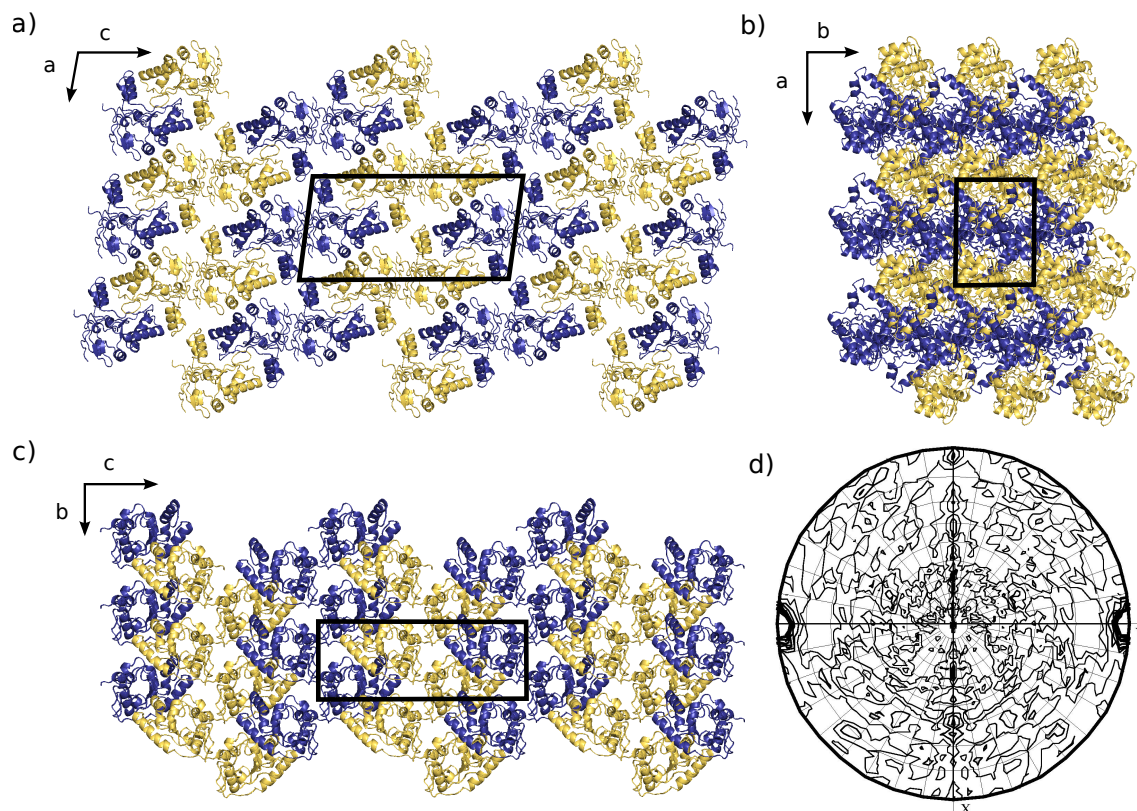
**Figura 5.30: Conformaciones abierta y cerrada de DacB<sub>49-238</sub>.** a) Superposición de las conformaciones abierta (morada) y cerrada (amarilla), mostrando el átomo de Zn<sup>2+</sup> en esferas y el fosfato en varillas. b) Detalle del lazo β<sub>5</sub> – α<sub>4</sub> en el que se observa un desplazamiento máximo de 16 Å, correspondiente a la Leu168. Otro cambio significativo es el producido en la orientación del Trp206, que en la conformación abierta se aleja 6 Å del sitio de unión a zinc (esfera y varillas) respecto a la conformación cerrada.

Como es habitual, por convenio, el eje binario helicoidal se corresponde con el eje *b*, de 37,6 Å de longitud, mientras que el ángulo β presenta un valor de 97,9°.

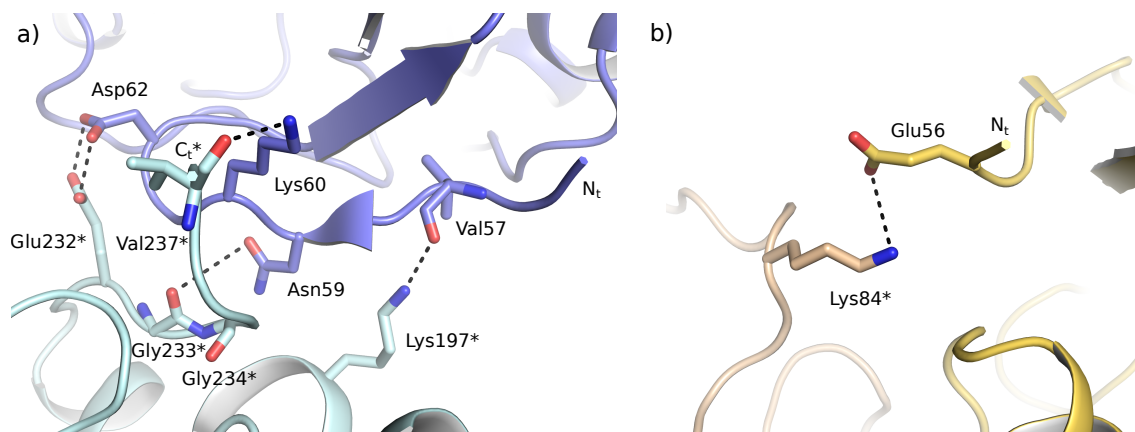
El empaquetamiento de DacB<sub>49-238</sub> incluye gran número de contactos no sólo entre las moléculas de la unidad asimétrica, sino también entre estas y las vecinas. No existe una orientación clara en la que los contactos sean mayoritarios, sino que toda la estructura contribuye a la interacción con las vecinas, generando un cristal compacto y sin apenas canales de solvente (Figura 5.31).

Mención especial merecen los primeros 6 residuos visibles de DacB<sub>49-238</sub> (AEVVNK, posiciones 55 a 60), ausentes en la construcción de DacB<sub>61-238</sub>. Como se puede observar en la Figura 5.32, esta región establece interacciones con las moléculas vecinas, influyendo significativamente en la orientación del empaquetamiento. La cadena cerrada de DacB<sub>49-238</sub> interacciona con otra cadena cerrada vecina, estabilizando sus extremos N y C-terminales respectivamente (Figura 5.32a), mientras que el extremo N-terminal de la cadena abierta establece un contacto con una zona intermedia de otra cadena abierta (Figura 5.32b). Esta pequeña diferencia de secuencia es la que ha permitido modificar la morfología de los cristales y evitar los problemas de macla característicos de DacB<sub>61-238</sub>.

Es interesante destacar la similitud del empaquetamiento de DacB<sub>49-238</sub> con el de las dos moléculas visibles de DacB<sub>20-238</sub> (Sección 5.2.4.2). Ambas construcciones se empaquetan en capas formadas por repeticiones de dos cadenas, una abierta y una cerrada, con sus sitios de unión a zinc enfrentados (Figura 5.33a). Además, las moléculas con la misma conformación (abierta o cerrada) enfrentan sus extremos N-terminales. La única diferencia existente entre los cristales de DacB<sub>49-238</sub> y DacB<sub>20-238</sub> es la separación entre capas. En los cristales de DacB<sub>49-238</sub> únicamente hay dos moléculas por unidad asimétrica (Figura 5.33b), mientras que, como se observó en el caso de DacB<sub>20-238</sub>, existe una tercera molécula que ocuparía el espacio libre entre capas (Figura 5.33c). Desde el punto de vista de la celdilla unidad, ambos cristales presentan dos

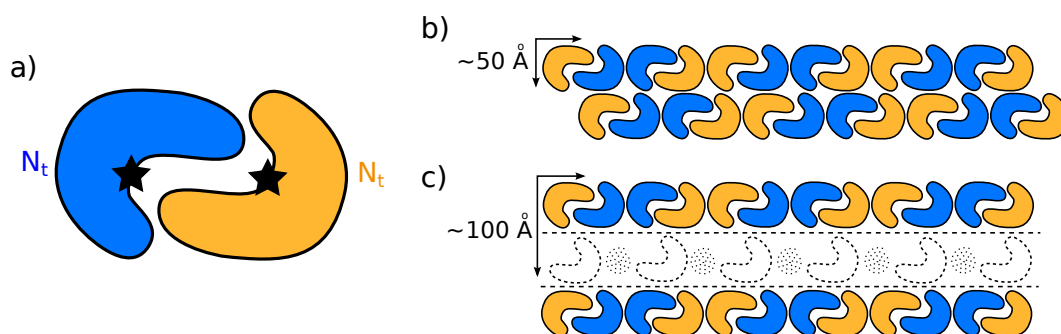


**Figura 5.31: Simetría y empaquetamiento de DacB<sub>49–238</sub>.** La conformación abierta se representa en morado, mientras que la conformación cerrada se representa en amarillo. El recuadro negro hace referencia a la celda unidad, indicándose además la orientación de cada eje en cada caso. **a)** Plano (*a, c*). **b)** Plano (*a, b*). **c)** Plano (*b, c*). **d)** Función de autorrotación calculada para  $\chi = 180^\circ$  mediante Molrep, mostrando los picos característicos de un grupo puntual  $2/m$ .



**Figura 5.32: Interacciones del extremo N-terminal de DacB<sub>49–238</sub>.** **a)** El extremo N-terminal de la cadena con conformación abierta (morado) interactúa con el extremo C-terminal de una cadena vecina, también de conformación abierta (azul). Los residuos Val57, Asn59 y Lys60, inexistentes en DacB<sub>61–238</sub>, establecen en este caso interacciones fuertes con la cadena lateral de Lys197 y las cadenas principales de Gly233 y Val237 respectivamente, todos ellos pertenecientes a la molécula vecina (etiquetados con \*). En consecuencia, surgen otras interacciones importantes para estabilizar el empaquetamiento, como la existente entre el Asp62 y la Glu232 vecina (Glu232\*). **b)** El extremo N-terminal de la cadena abierta de DacB<sub>49–238</sub> (amarillo) establece un único contacto con una cadena vecina (rosa), también abierta, mediado por la Glu56, inexistente en DacB<sub>61–238</sub>. En ambos paneles, los residuos etiquetados con \* pertenecen a una molécula vecina.

ejes de aproximadamente 37 y 103 Å, mientras que el tercer eje en los cristales de DacB<sub>20–238</sub> es sensiblemente más largo que en el caso de DacB<sub>49–238</sub> (aproximadamente el doble). Esto se correlaciona con el hecho de que la distancia entre capas de los cristales de DacB<sub>20–238</sub> equivale al tamaño de una única cadena, es decir, a una capa adicional de ancho equivalente a la capa observable (Figura 5.33c). Por último, es importante destacar que las condiciones de cristalización de ambas construcciones son altamente similares. Como se indicó en la Sección 5.4.2, tanto DacB<sub>20–238</sub> como DacB<sub>49–238</sub> cristalizaron a la misma concentración de ZnCl<sub>2</sub>, a concentración similar de PEG 6000 y a pH 8,0. Incluso el hábito cristalino se mantuvo, siendo de mayor calidad en el caso de DacB<sub>49–238</sub>. La única diferencia fue la presencia y ausencia del detergente CYMAL-3 en DacB<sub>20–238</sub> y DacB<sub>49–238</sub> respectivamente. Así pues, en base al análisis del empaquetamiento y a la comparación del proceso de cristalización, se reafirma la hipótesis de que el espacio vacío de los cristales de DacB<sub>20–238</sub> está ocupado por una tercera molécula y los fragmentos N-terminales desordenados de las tres cadenas, parcialmente estabilizados con ayuda del CYMAL-3. Los cristales de DacB<sub>49–238</sub>, al no contener el fragmento N-terminal desordenado, no necesitan CYMAL-3 para estabilizar dicha región (es más, la presencia de CYMAL-3 no sólo no es necesaria sino que impide la cristalización de esta construcción).



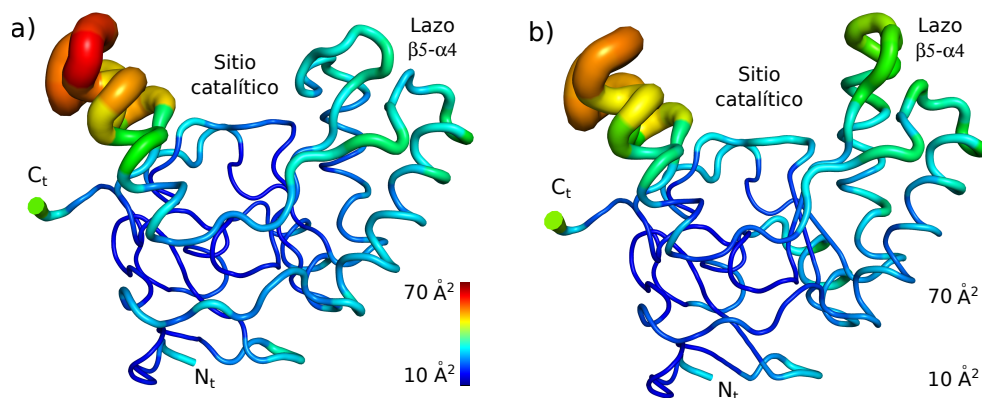
**Figura 5.33: Comparación del empaquetamiento de DacB<sub>20–238</sub> y DacB<sub>49–238</sub>.** **a)** En ambos cristales, cadenas abiertas y cerradas (azul y naranja respectivamente) se enfrentan entre sí. Los sitios de unión a zinc se representan mediante una estrella. Además, se indican las zonas correspondientes al extremo N-terminal de cada cadena (N<sub>t</sub>). **b)** Esquema del empaquetamiento de DacB<sub>49–238</sub>, mostrando el ancho aproximado de cada capa que compone la totalidad del cristal. **c)** Esquema del empaquetamiento de DacB<sub>20–238</sub>, mostrando que el ancho de cada unidad de repetición (correspondiente a una capa más un espacio vacío), es aproximadamente el doble que en el caso de DacB<sub>49–238</sub>. En puntos se representan las posiciones teóricas de la tercera cadena (ausente en el modelo) y de los extremos N-terminales desordenados (nube de puntos).

#### 5.4.4.3. Agitación térmica

Las cadenas cerrada y abierta de DacB<sub>61–238</sub> presentan unos factores B medios bajos (23,4 y 24,0 Å<sup>2</sup> respectivamente), por lo que en general la estructura es estable. No obstante, existe un lazo expuesto (Asp131 – Tyr144) que presenta factores de agitación térmica superiores a la media, indicando una mayor flexibilidad respecto al resto de la estructura. En la conformación abierta de DacB<sub>49–238</sub>, los factores B para la cadena principal de este lazo flexible alcanzan valores de 70 Å<sup>2</sup> (Figura 5.34a), mientras que en la conformación cerrada no superan los 60 Å<sup>2</sup> (Figura 5.34b). La agitación de esta zona se debe a su orientación hacia una zona expuesta, siendo la única región de DacB<sub>49–238</sub> que carece de estabilización cercana con las moléculas vecinas.



El lazo  $\beta_5 - \alpha_4$  presenta factores B bajos pero diferentes entre ambas estructuras. Mientras que en la estructura cerrada se sitúan en torno a  $20 \text{ \AA}^2$  (Figura 5.34a), en la estructura abierta alcanzan los  $40 \text{ \AA}^2$  (Figura 5.34b). Ya que los factores B medios globales de ambas estructuras coinciden, la diferencia en la agitación del lazo  $\beta_5 - \alpha_4$  es significativa. No obstante, mientras que el lazo  $\beta_5 - \alpha_4$  de la estructura cerrada presenta interacciones con la molécula abierta de la unidad asimétrica, el lazo  $\beta_5 - \alpha_4$  de la estructura abierta se estabiliza mediante contactos con una cadena cerrada vecina. Así pues, la diferencia entre los factores B del lazo  $\beta_5 - \alpha_4$  puede deberse más a la organización del empaquetamiento que a una causa fisiológica de la enzima.



**Figura 5.34: Representación de factores de agitación térmica de DacB<sub>49-238</sub>.** Los valores cercanos al azul reflejan baja agitación térmica, mientras que un virado hacia el rojo implica mayor agitación térmica y mayor flexibilidad. Para mayor claridad, se han indicado la posición del sitio catalítico y del lazo  $\beta_5 - \alpha_4$ . **a)** Conformación cerrada. **b)** Conformación abierta.

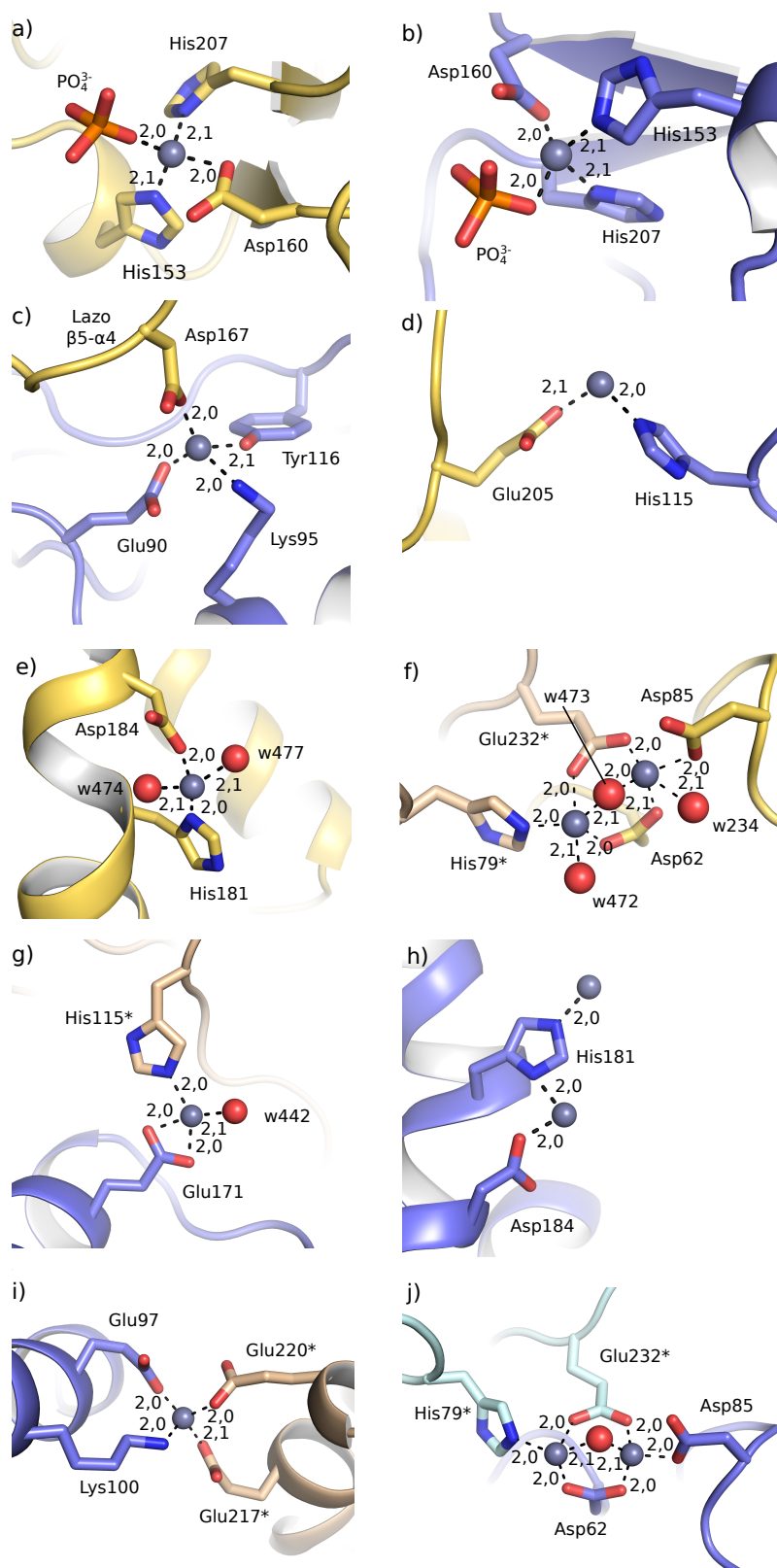
#### 5.4.4.4. Átomos de zinc y solvente

Como ya se ha indicado, se encontró un  $\text{Zn}^{2+}$  en cada sitio de unión a zinc, coordinado con la tríada His153 – Asp160 – His207. La coordinación del  $\text{Zn}^{2+}$  se completa con una molécula de fosfato, también presente en ambas cadenas (Figuras 5.35a y 5.35b). Los factores B del  $\text{Zn}^{2+}$  son  $15,0 \text{ \AA}^2$  para la conformación cerrada y  $17,5 \text{ \AA}^2$  para la conformación abierta, mientras que los fosfatos presentan valores en torno a  $30$  y  $40 \text{ \AA}^2$  respectivamente.

Se encontraron además 11 átomos de  $\text{Zn}^{2+}$  adicionales repartidos alrededor de las dos cadenas de la unidad asimétrica, favoreciendo contactos tanto entre ellas como con las moléculas vecinas. Si bien algunos favorecían el empaquetamiento entre las dos moléculas de la unidad asimétrica (Figuras 5.35c y 5.35d), otros aparecían alrededor de la estructura, bien coordinándose con alguna molécula de agua del solvente o con residuos de cadenas vecinas (Figuras 5.35e a 5.35j), favoreciendo aún más el empaquetamiento del cristal.

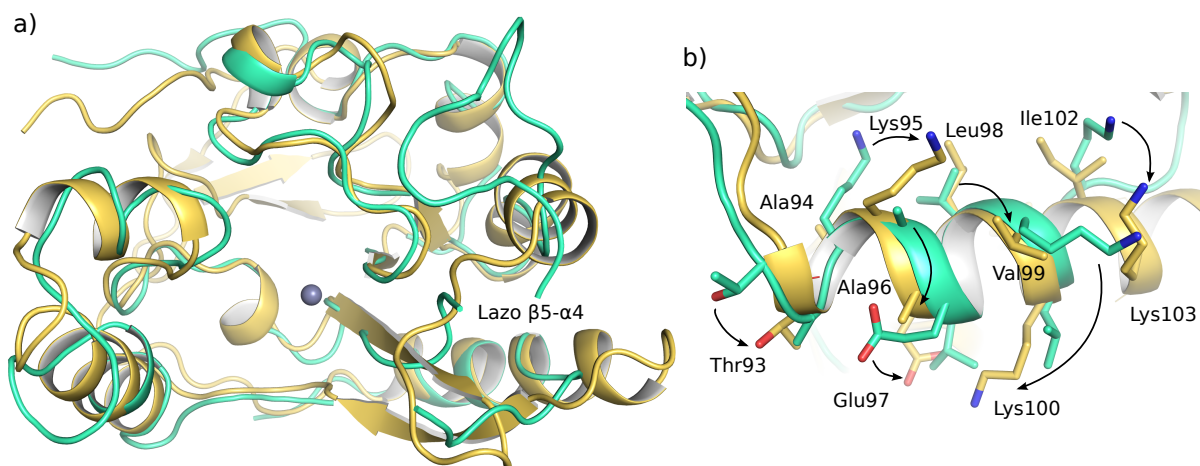
#### 5.4.5. Relación entre DacB<sub>49-238</sub> y DacB<sub>61-238</sub>

Una vez obtenida la estructura de DacB<sub>49-238</sub>, se pueden intentar resolver los datos maclados de los cristales de DacB<sub>61-238</sub>. Es de esperar que el modelo de DacB<sub>49-238</sub>, si bien no idéntico, debe presentar una alta similitud estructural con el modelo de DacB<sub>61-238</sub>. Así pues, se dispone de un mejor modelo que el de DacB<sub>20-238</sub>, utilizado originalmente para resolver las maclas.



**Figura 5.35: Interacciones de átomos de zinc en DacB<sub>49-238</sub>.** Las cadenas abierta y cerrada de la celdilla unidad se representan en morado y amarillo respectivamente. Las cadenas abiertas y cerradas vecinas se representan en azul y marrón respectivamente, estando además sus residuos etiquetados con \*. Las moléculas de agua se representan en esferas rojas y los átomos de zinc en esferas moradas. Todas las distancias vienen dadas en Å. **a)** Sitio catalítico de la cadena cerrada. **b)** Sitio catalítico de la cadena abierta. **c y d)** Átomos puente entre las cadenas de la unidad asimétrica. **e y f)** Átomos de zinc propios de la cadena cerrada o compartidos con moléculas vecinas. **g, h, i y j)** Átomos de zinc propios de la cadena abierta o compartidos con moléculas vecinas.

Los mapas obtenidos al realizar un reemplazo molecular de DacB<sub>49–238</sub> sobre los datos de DacB<sub>61–238</sub> son sensiblemente mejores que los conseguidos durante el procesamiento de los cristales de DacB<sub>61–238</sub>, lo que sugiere la existencia de diferencias sustanciales entre ambos modelos. De este modo, se puede comprobar que algunos elementos de estructura secundaria de DacB<sub>61–238</sub> se encontraban desplazados respecto a los reales (DacB<sub>49–238</sub>), inapreciables en aquel momento debido al sesgo de los mapas (Figura 5.36).



**Figura 5.36: Comparación estructural entre DacB<sub>49–238</sub> y DacB<sub>61–238</sub>.** La estructura de DacB<sub>49–238</sub> se representa en amarillo. La estructura de DacB<sub>61–238</sub> se representa en azul. **a)** Superposición general en la que se observan las discrepancias existentes a nivel de cadena principal. **b)** Detalle de la hélice  $\alpha_1$  en el que se aprecia el desfase en las posiciones de las cadenas laterales, frecuente en varios elementos de estructura secundaria.

No obstante, al enfrentar el modelo de DacB<sub>49–238</sub> con los datos de difracción de DacB<sub>61–238</sub>, los factores  $R$  de refinamiento se mantienen en torno a 0,30 – 0,32, sin posibilidad de mejoría sustancial. Este hecho demuestra que no sólo un modelo imperfecto impidió la resolución de DacB<sub>61–238</sub>, sino que, además, los datos de difracción no alcanzaban la calidad suficiente como para resolver la estructura con una mínima garantía estadística.

#### 5.4.6. Cristalización en presencia de ligandos

Para completar el estudio de DacB<sub>49–238</sub> y elucidar su mecanismo de acción, se realizaron experimentos tanto de inmersión como de cocrystalización en presencia de ligandos, siguiendo las técnicas descritas en la Sección 3.2.3. Como ligandos se utilizaron cuatro fragmentos de peptidoglicano de origen sintético, producidos por el grupo del Prof. Shahriar Mobashery (Departamento de Química y Bioquímica, Universidad de Notre Dame, Indiana, EEUU).

Como se introdujo en la Sección 1.3.3, la proteína DacB cataliza la conversión de tetrapéptido (L-Ala–D-iGln–L-Lys–D-Ala) a tripéptido (L-Ala–D-iGln–L-Lys) tras la eliminación de la D-Ala terminal. Así pues, es de esperar que el sustrato sea bien el tetrapéptido aislado, bien el tetrapéptido junto con la parte glucídica perteneciente a la estructura principal del peptidoglicano (GlcNAc–MurNAc). Por tanto, se intentó la cristalización del complejo de DacB con el tetrapéptido aislado y con el tetrapéptido unido a MurNAc. Además, ya que los productos de reacción también están en contacto directo con la enzima hasta su liberación, se ensayaron igualmente el tripéptido aislado y el tripéptido unido a MurNAc.

Se realizaron una serie de experimentos de cocrystalización incubando la proteína con cada ligando por separado a diferentes concentraciones y diferentes tiempos. En general, se ensayó un rango de 0,1 mM a 50 mM para cada ligando, con tiempos de incubación de entre 30 minutos y 24 horas. Los ensayos de inmersión se realizaron en el mismo rango de concentraciones de cada ligando, con tiempos que oscilaban entre pocos segundos y 5 minutos. En ningún caso se consiguió obtener un complejo entre la proteína y el ligando.

Los ensayos de cocrystalización siempre devolvieron cristales de la proteína libre, mientras que en los ensayos de inmersión los cristales se deterioraban con rapidez. De este comportamiento se extrae que los sustratos sintéticos, por ser idénticos a los naturales, eran procesados por la enzima en el cristal, transformados en sus respectivos productos y liberados al medio. Siguiendo el mismo razonamiento pero para los productos, si estos llegan a entrar en el sitio catalítico serán inmediatamente liberados al medio, por lo que seguirá sin formarse complejo con DacB.

A la vista de estos resultados se intentó eliminar el  $\text{Zn}^{2+}$  de los cristales de DacB<sub>49–238</sub> para obtener la apoenzima y, posteriormente, realizar inmersiones en los distintos ligandos sintéticos, con la intención de que los sustratos no fuesen catalizados y permaneciesen unidos al sitio catalítico. Para ello, se sometió a los cristales a un tratamiento de inmersión en EDTA a diferentes concentraciones y diferentes tiempos. En general, los cristales se disolvían tras entrar en contacto con el EDTA, algo esperable dado el gran número de átomos de zinc que participan en el empaquetamiento. Los que no se disolvían, bien mantenían el  $\text{Zn}^{2+}$  en el sitio activo, bien sufrían daños visualmente imperceptibles pero que alteraban la difracción hasta convertirlos en irresolubles. Así pues, esta estrategia tampoco reportó resultados positivos en la búsqueda del complejo DacB : ligando.

## 5.5. Relación estructura-función

Como se indicó en la Sección 1.3.3, la degradación del peptidoglicano es una función de vital importancia para las bacterias Gram positivas. Varias proteínas se dedican a catalizar la rotura de este componente de la pared bacteriana en diversos fragmentos, siendo de especial interés las proteínas DacA (D,D-carboxipeptidasa) y DacB (L,D-carboxipeptidasa), que trabajan en tándem y catalizan la conversión de pentapéptidos a tetrapéptidos y de estos últimos a tripéptidos, respectivamente. Los resultados estructurales presentados en este trabajo, junto con experimentos microbiológicos y fisiológicos han permitido ampliar el conocimiento existente sobre las proteínas DacA, DacB y su mecanismo de acción.

### 5.5.1. Maduración de DacB

DacB es una lipoproteína que se sintetiza en forma de pre-prolipoproteína con un péptido señal asociado al extremo N-terminal. El péptido señal contiene, en su extremo C-terminal, el motivo *LipoBox*, que en el caso de DacB se corresponde con Leu16–Ala17–Ala18–Cys19. En primer lugar, la pre-prolipoproteína diacilgliceril transferasa Lgt establece un enlace covalente entre la Cys19 del *LipoBox* y el fragmento de diacilglicerol de la membrana lipídica. Posteriormente, la peptidasa de señal LspA hidroliza el enlace entre la Ala18 y la Cys19, retirando



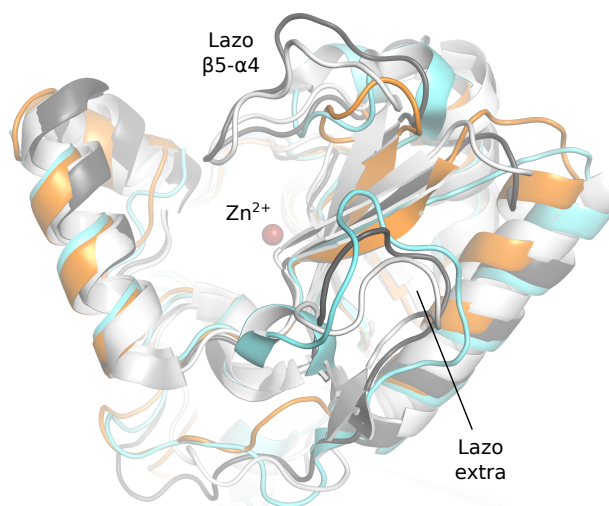
el péptido señal de la estructura. De esta forma se obtiene la lipoproteína DacB madura anclada covalentemente a la membrana (Rezwan *et al.*, 2007; Kovacs-Simon *et al.*, 2011).

Los predictores de estructura secundaria (TMpred y PSIPRED; Hofmann & Stoffel, 1993; Buchan *et al.*, 2013) indican que el péptido señal podría presentar un plegamiento de hélice transmembrana, quedando el motivo *LipoBox* alineado con la capa externa de los fragmentos de diacilglicerol. Así pues, cabe la posibilidad de que la pre-prolipoproteína DacB utilice el péptido señal a modo de hélice transmembrana para insertarse en la membrana y translocarse hacia el exterior. Posteriormente tendría lugar la maduración de la proteína, eliminando el péptido señal transmembrana y estableciendo el enlace covalente entre la Cys19 y el diacilglicerol.

### 5.5.2. Análisis estructural comparado

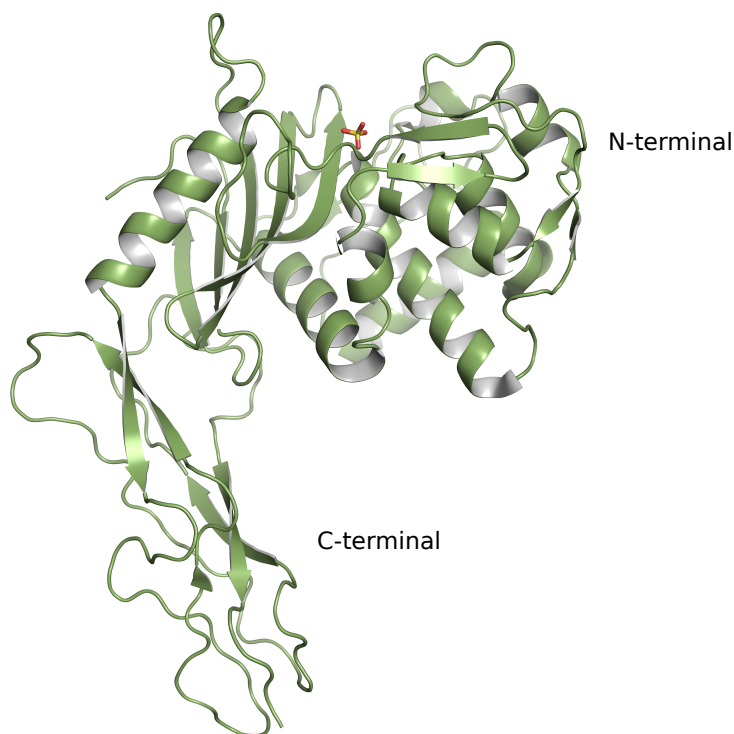
La búsqueda de estructuras similares a DacB, realizada mediante el servidor Dali (Holm & Rosenström, 2010) devolvió la proteína BaLdcB de *B. anthracis* (VanY-like L-lisil-D-alanina carboxipeptidasa, código PDB 4MPH; Hoyland *et al.*, 2014) como la de mayor similitud resuelta hasta la fecha, con un RMSD de 1,8 Å para 184 C $\alpha$ . Otras proteínas relacionadas mostraron una similitud significativa, como las D-alanil-D-alanina carboxipeptidasas VanXYc de *E. gallinarum* (código PDB 4MUS, RMSD = 2,1 Å para 188 C $\alpha$ ) y VanXYg de *E. faecalis* (código PDB 4F78, RMSD = 2,4 Å para 251 C $\alpha$ ; Meziane-Cherif *et al.*, 2014).

En general, tanto el plegamiento como el sitio catalítico presentan un alto grado de conservación entre estas proteínas y DacB. No obstante, existen dos regiones en las que las diferencias son significativas (Figura 5.37). En primer lugar, el lazo  $\beta_5 - \alpha_4$  es más corto en DacB que en VanXYc/g y no llega a cubrir el sitio catalítico al completo, sucediendo lo mismo con el lazo de BaLdcB. En segundo lugar, DacB carece de un lazo de 20 aminoácidos que tanto BaLdcB como VanXYc/g presentan antes de la hélice  $\alpha_1$ , cerrando el sitio catalítico.



**Figura 5.37: Superposición estructural de DacB con proteínas similares.** La conformación cerrada de DacB se representa en naranja, mientras que BaLdcB aparece en azul, VanXYc en gris y VanXYg en blanco. El lazo  $\beta_5 - \alpha_4$  de DacB presenta un tamaño y orientación similar al lazo equivalente de BaLdcB, mientras que los respectivos lazos de las proteínas VanXYc/g, de longitud mayor, se sitúan sobre el sitio de unión a zinc. Además, DacB carece del lazo extra que flanquea el sitio catalítico en el resto de proteínas.

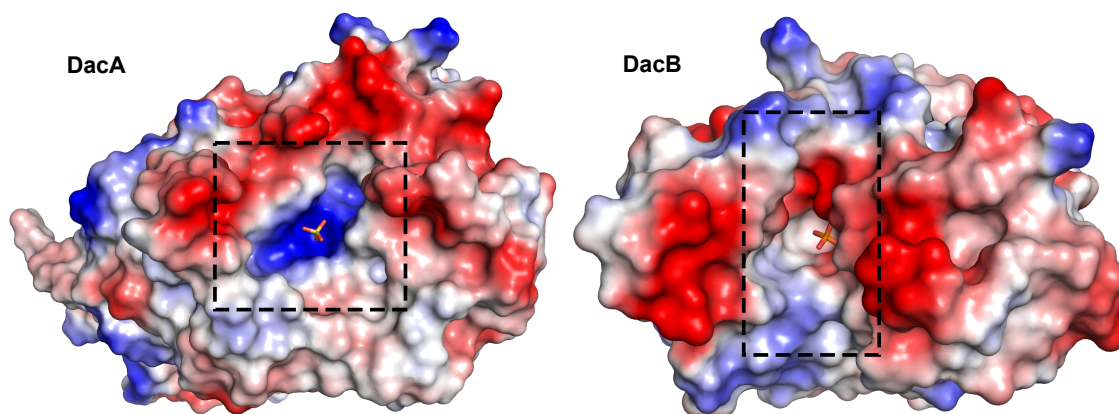
Al comparar DacB con la proteína DacA de *S. pneumoniae* (código PDB 1XP4; Morlot *et al.*, 2005), se observa un plegamiento y organización completamente diferente entre sí. Mientras que el dominio catalítico de DacB sigue un plegamiento clásico de la superfamilia de las peniciloil serina transferasas, en el que 5 hebras  $\beta$  antiparalelas ocupan una región central, rodeadas de hélices  $\alpha$ , DacA contiene un dominio de tipo D,D-carboxipeptidasa N-terminal y una región C-terminal rica en hebras  $\beta$  (Figura 5.38).



**Figura 5.38: Estructura general de DacA.** La estructura se divide en dos dominios, a saber, el N-terminal, con plegamiento típico de D,D-carboxipeptidasas, y el C-terminal, rico en hebras  $\beta$ .

Las  $\beta$ -lactamasas de clase A contienen el denominado lazo  $\Omega$ , que incluye residuos implicados tanto en el mantenimiento de la estructura del centro activo como en la actividad enzimática de la proteína. Este hecho es de especial interés ya que las D,D-carboxipeptidasas pertenecientes a bacterias Gram positivas presentan un lazo  $\Omega$  de 14 aminoácidos, factor diferencial con respecto a las D,D-carboxipeptidasas de bacterias Gram negativas (Morlot *et al.*, 2005). La proteína DacB de *S. pneumoniae* no contiene remanentes del lazo  $\Omega$ , lo que encaja con su actividad L,D-carboxipeptidasa y no D,D-carboxipeptidasa.

El potencial electrostático de la superficie de DacB también difiere del de DacA. La proteína DacA presenta un sitio activo estrecho de fuerte carácter básico, mientras que DacB contiene una cavidad alargada principalmente ácida (Figura 5.39). Este hecho encaja con el reconocimiento de sustrato que realizan ambas proteínas y favorece su especificidad, ya que, mientras que la proteína DacB necesita una superficie ácida afín por la L-Lys del tetrapéptido, la proteína DacA evita la entrada de este residuo utilizando una superficie de carácter básico.



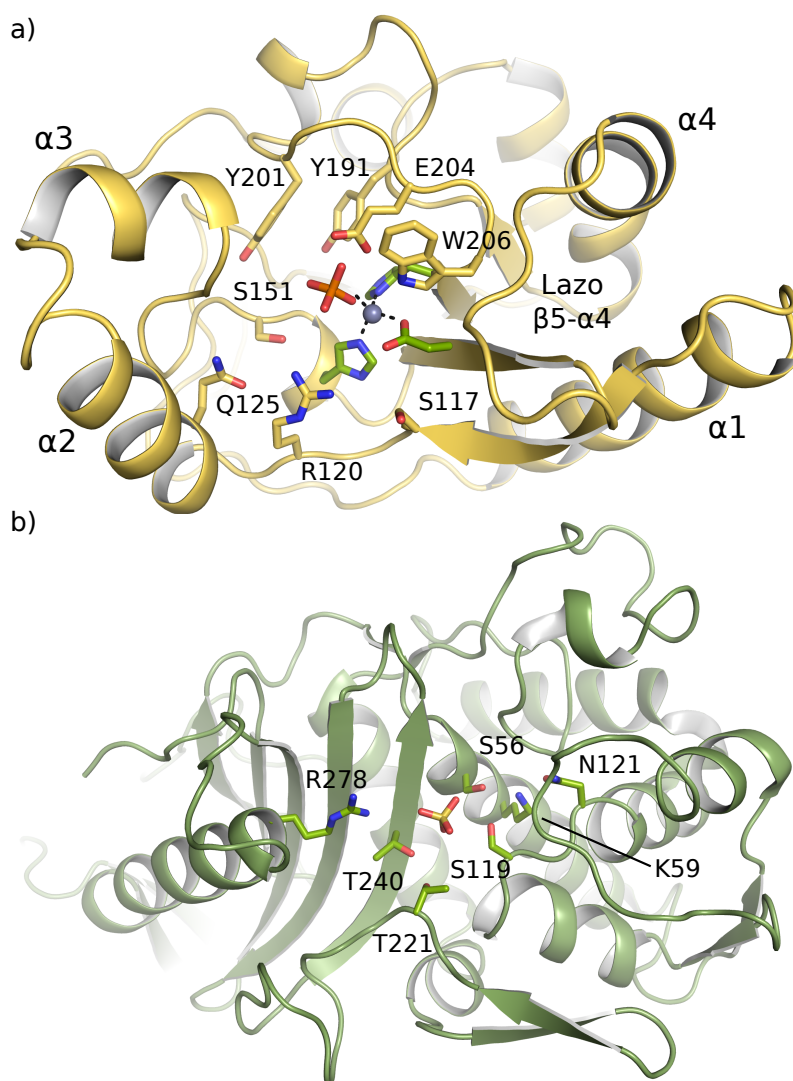
**Figura 5.39: Potencial electrostático de las superficies de DacA y DacB.** En azul se representan las regiones básicas, mientras que en rojo aparecen las regiones ácidas. Los recuadros indican los sitios de unión a sustrato en cada proteína. En varillas se han representado las moléculas de sulfato y fosfato observadas en los sitios catalíticos de DacA y DacB respectivamente.

### 5.5.3. Maquinaria catalítica y conformaciones estructurales

Como se vio en la Sección 5.4.4.4, el eje central de la maquinaria catalítica de DacB es el sitio de unión a zinc, formado mediante coordinación tetrahédrica con dos histidinas (His153 e His207), un ácido aspártico (Asp160) y un radical hidroxilo de una molécula de fosfato del solvente. Por analogía con las metaloproteasas dependientes de zinc VanXYc/g (Meziane-Cherif *et al.*, 2014), es probable que el residuo que actúe como catalizador ácido/base sea el Glu204, que ya presenta interacciones con el fosfato que coordina el  $\text{Zn}^{2+}$ . Así pues, el zinc y el Glu204 actuarían conjuntamente promoviendo el ataque de una molécula de agua sobre el carbono del grupo carbonilo del enlace peptídico del sustrato. Posteriormente, el Glu204 enviaría dos protones al nitrógeno del enlace peptídico, favoreciendo su rotura. La estabilización de la carga negativa en el intermediario tetrahédrico formado durante el estado de transición se podría llevar a cabo mediante la Arg120, como se ha observado en enzimas relacionadas (Korndörfer *et al.*, 2008), mientras que el nitrógeno doblemente protonado podría verse estabilizado mediante la acción, por proximidad, de la Tyr191 o de la Tyr201. Además, la Arg120 presenta una orientación estable mediante interacción de enlace de hidrógeno con la Gln125 (Figura 5.40a).

Este mecanismo de reacción es muy distinto al realizado por DacA, la cual, si bien ejerce una actividad carboxipeptidasa, presenta una maquinaria catalítica formada por tres motivos estructurales conservados, a saber, SXXK (Ser56 – Ile57 – Thr58 – Lys59, siendo la Ser56 el residuo nucleofílico), SXN (Ser119 – Ala120 – Asn121) y K(T/S)G (Lys239 – Thr240 – Gly241) (Morlot *et al.*, 2005) (Figura 5.40b).

Como se vio en la Figura 5.28, el sitio de unión a zinc de DacB y los residuos accesorios que participan en la reacción se ubican en una cavidad formada sobre la lámina  $\beta$  central, flanqueada por las hélices  $\alpha_2$  y  $\alpha_3$  por un lado y  $\alpha_1$  y  $\alpha_4$  por el otro. La conformación abierta de DacB presenta la cavidad más grande, con unas dimensiones aproximadas de  $24 \times 15 \text{ \AA}$  (Figura 5.41a), mientras que la conformación cerrada reduce la cavidad hasta unas dimensiones aproximadas de  $24 \times 10 \text{ \AA}$  (Figura 5.41b). En el caso de VanXYc, el lazo se sitúa sobre el sitio catalítico,

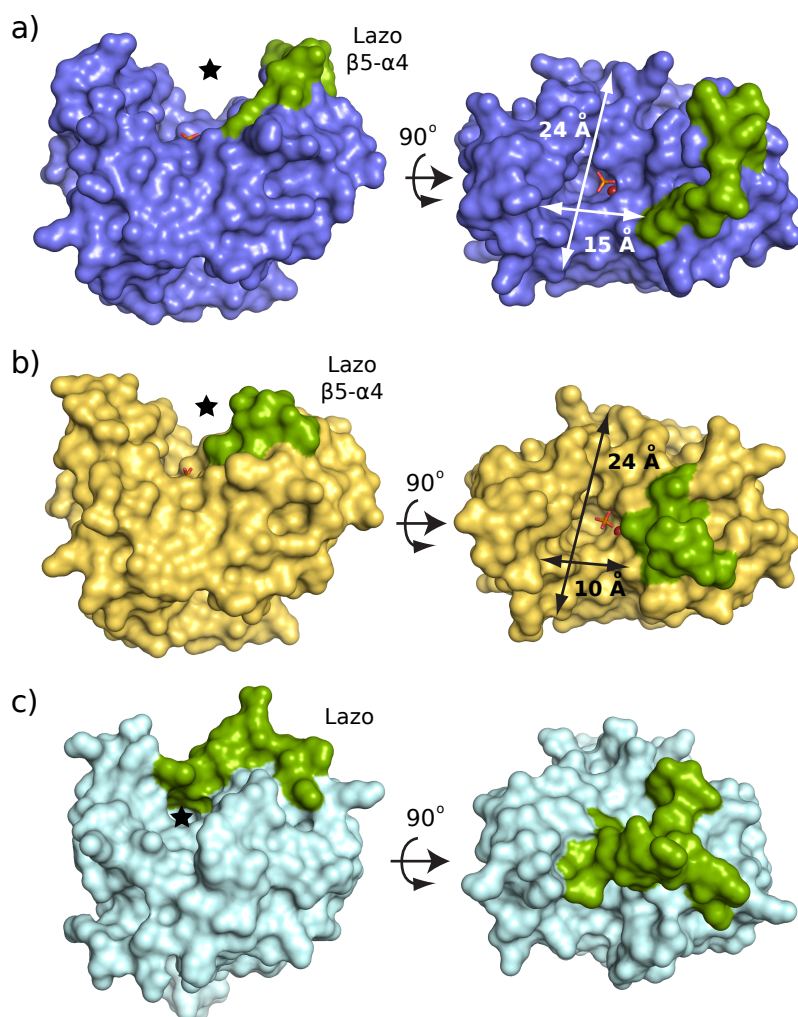


**Figura 5.40: Sitios catalíticos de DacB y DacA.** a) Sitio catalítico de DacB en conformación cerrada. Se muestran en varillas los residuos que, por su cercanía al sitio de unión a zinc, pueden estar implicados en la catálisis o estabilización del sustrato. En verde se muestran los residuos que coordinan al  $\text{Zn}^{2+}$ . b) Sitio catalítico de DacA. En varillas se muestran los aminoácidos más representativos.

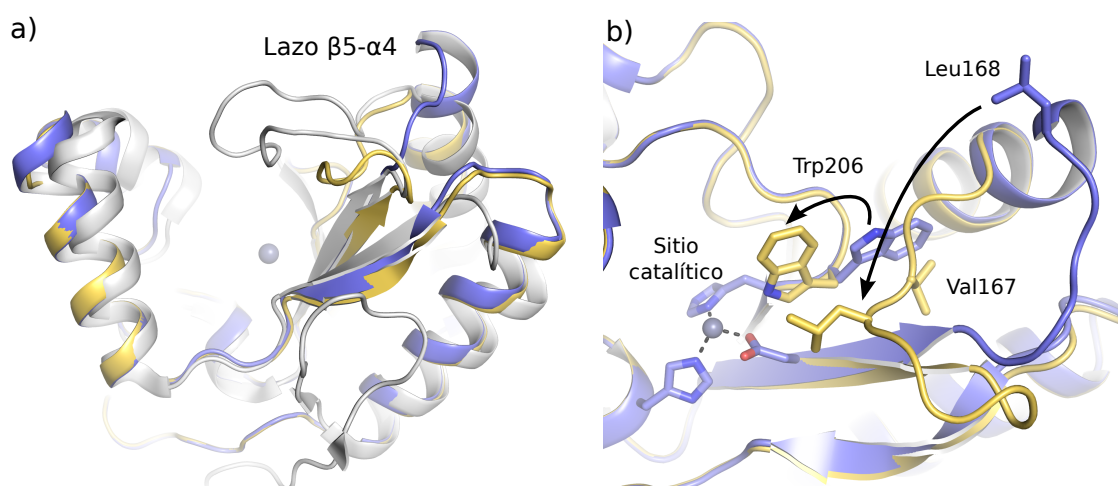
dejando un túnel estrecho por el que accede el sustrato en una orientación diferente a como lo hace en DacB (Figura 5.41c).

La comparación entre ambas cadenas de la unidad asimétrica de DacB<sub>49–238</sub> revela cambios no sólo en la orientación del lazo  $\beta_5 - \alpha_4$ , sino también en la estructura de la hélice  $\alpha_4$ , que sufre una reorganización al pasar de una conformación a otra (Figura 5.42a). La hélice  $\alpha_4$  de la conformación abierta es más larga que en el caso de la conformación cerrada (Leu168 a Tyr185 y Glu172 a Tyr185 respectivamente). Así pues, en el segundo caso, el lazo  $\beta_5 - \alpha_4$  aumenta el número de residuos y alcanza el centro activo. De esta forma, residuos como la Leu168 pasan de estar a una distancia aproximada de 20 Å del sitio de unión a zinc en la conformación abierta a acercarse hasta una distancia de 7 Å en la conformación cerrada. El Trp206 sufre un cambio drástico de orientación, abandonando un bolsillo hidrofóbico para acercarse al sitio catalítico, mientras que el hueco que ha dejado libre es ocupado por la Val167 (Figura 5.42b).





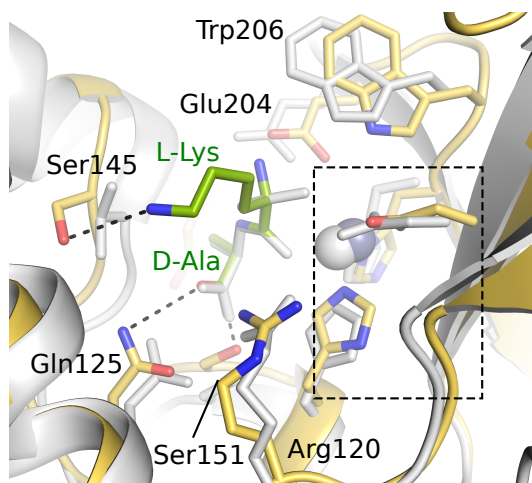
**Figura 5.41: Sitios de unión a sustrato en DacB y VanXYc.** El lazo  $\beta 5-\alpha 4$  o equivalente se representa en color verde. El sitio de unión a sustrato se representa mediante una estrella.  
**a)** Conformación abierta de DacB. **b)** Conformación cerrada de DacB. **c)** Estructura de VanXYc.



**Figura 5.42: Comparación de los lazos  $\beta 5-\alpha 4$  de las conformaciones abierta y cerrada de DacB.** En morado se representa la conformación abierta de DacB, mientras que en amarillo se representa la conformación cerrada. **a)** Se aprecia la diferente orientación del lazo entre las diferentes conformaciones de DacB. Se ha incluido la estructura de VanXYc (en blanco) como modelo de estructura con lazo largo. **b)** Detalle de los cambios estructurales relevantes debidos al cierre de la conformación de DacB.

En el caso de las D,D-carboxipeptidasas VanXYc y VanXYg, sus respectivos lazos (equivalentes al lazo  $\beta_5 - \alpha_4$  de DacB) cubren la totalidad del sitio activo, creando un túnel estrecho por el que se introduce el sustrato, D-Ala–D-Ala. La cristalización de estas proteínas en presencia de un análogo del estado de transición ha permitido determinar los residuos implicados en el reconocimiento del sustrato. Ya que la maquinaria catalítica de las D,D-carboxipeptidasas de *E. faecalis* y *E. gallinarum* es equivalente a la de la proteína DacB, se realizó un modelo computacional del sustrato L-Lys–D-Ala, utilizando como base el fragmento de D-Ala–D-Ala de las proteínas VanXYc/g. La orientación del fragmento de L-Lys–D-Ala se determinó mediante la superposición del complejo VanXYc : D-Ala–D-Ala y la conformación cerrada de DacB, de mayor similitud que la conformación abierta.

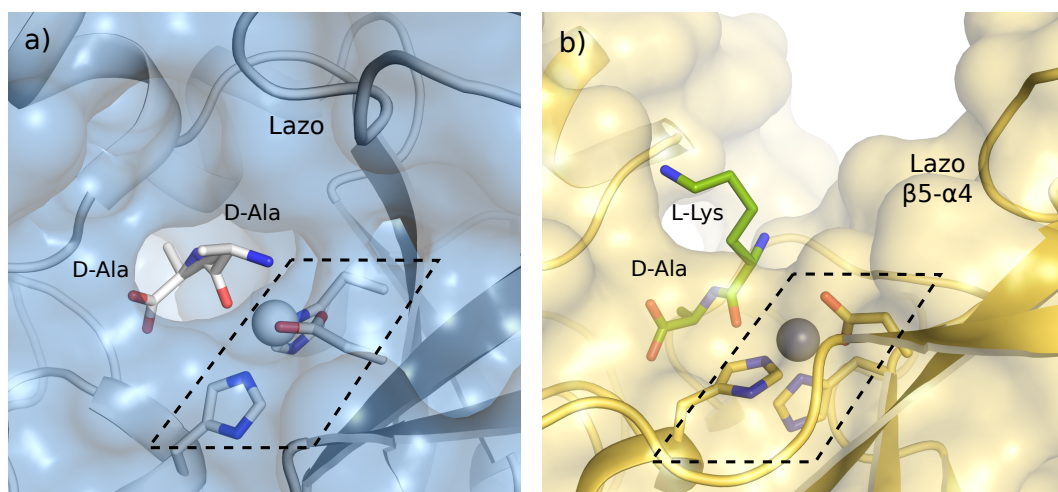
La estabilización de la D-Ala se realiza de manera similar en ambas proteínas, algo esperable ya que los sustratos naturales de ambas proteínas contienen D-Ala como residuo terminal. En cambio, aparecen diferencias notables en el sitio que rodea a la L-Lys en DacB, respecto al mismo sitio en VanXYc : D-Ala–D-Ala, ya que, en DacB, existe un hueco en el que se puede acomodar la cadena lateral de la L-Lys. La Ser145, ausente en VanXYc, podría ser la encargada de estabilizar a la L-Lys (Figura 5.43). Este hecho es de especial relevancia ya que, en VanXYc, esta posición se encuentra ocupada por la Val87, por lo que la existencia de la serina permite que aparezca en la zona un residuo polar con capacidad de establecer enlaces de hidrógeno. Los residuos implicados en la estabilización de la cadena principal del dipéptido están conservados entre ambas estructuras. En concreto, la Gln125 y la Ser151 podrían interactuar con el extremo C-terminal de la D-Ala, mientras que el Glu204 catalítico y el Trp206 podrían estabilizar el resto del sustrato (Figura 5.43).



**Figura 5.43: Modelo computacional de L-Lys–D-Ala y residuos implicados en su reconocimiento.** La conformación cerrada de DacB se representa en amarillo, superpuesta a la estructura de VanXYc, en blanco. El modelo computacional del péptido L-Lys–D-Ala se muestra en verde, superpuesto al fragmento D-Ala–D-Ala de VanXYc, en blanco. El sitio de unión a zinc se ha recuadrado con líneas punteadas. Los átomos de zinc de DacB y VanXYc se representan en esferas roja y blanca respectivamente.

Es interesante también analizar la orientación de la cadena principal del modelo de L-Lys–D-Ala, ya que difiere en quiralidad respecto a la cadena de D-Ala–D-Ala de VanXYc. Mientras que en VanXYc la única orientación posible para el resto del sustrato debe hacerse a

través del túnel formado entre el sitio catalítico y el lazo equivalente al lazo  $\beta_5 - \alpha_4$ , en DacB debe ubicarse en la única zona de la cavidad expuesta al solvente (Figura 5.44). Por ello, es probable que el lazo  $\beta_5 - \alpha_4$  no sólo prepare cavidad catalítica de DacB para que se produzca la reacción sino que esté directamente implicado en el reconocimiento y orientación del sustrato, aportando algún residuo capaz de interactuar con el D-iGln o la L-Ala del fragmento restante del tetrapéptido. Residuos como el Asp167 o la Thr170 (Figura 5.30) pueden ser candidatos para estabilizar las cadenas de peptidoglicano mientras se produce el corte del tetrapéptido.



**Figura 5.44: Reconocimiento de sustrato en VanXYc y DacB.** a) En VanXYc, el túnel estrecho formado por el lazo equivalente al lazo  $\beta_5 - \alpha_4$  limita la orientación del sustrato restante en una única dirección. b) El lazo  $\beta_5 - \alpha_4$  de DacB, por ser más corto, no limita la orientación del fragmento restante del sustrato, pudiendo ubicarse sobre la superficie expuesta al solvente con relativa libertad de movimiento. En ambos paneles, el sitio de unión a zinc se ha recuadrado como referencia.

Por último, la presencia en DacB de una región desordenada de unos 30 aminoácidos entre el anclaje a membrana y el inicio del dominio catalítico sugiere que esta proteína presenta una flexibilidad característica de la que carece DacA (Figura 5.38). Este hecho sugiere que ambas proteínas podrían realizar su función en diferentes regiones de la pared, lo que supondría una regulación en el proceso de degradación de peptidoglicano. Además, el hecho de que DacB contenga el lazo  $\beta_5 - \alpha_4$  conlleva la existencia de un mecanismo de regulación de su actividad mediante cambios conformacionales, mientras que DacA se mantiene constantemente activa. Debido a ello, no se descarta la posible existencia de un complejo funcional de DacA o DacB con otras proteínas implicadas en la remodelación del peptidoglicano, regulándose así la tasa y localización de su degradación.

#### 5.5.4. Estudios microbiológicos y fisiológicos relacionados con DacA y DacB

Los experimentos descritos a lo largo de esta sección han sido realizados por el grupo del Prof. Sven Hammerschmidt (Departamento de Genética Microbiana de la Universidad de Greifswald, Alemania). No obstante, es necesario realizar una visión general de los resultados obtenidos ya que se encuentran estrechamente relacionados con la estructura de DacB obtenida en este trabajo y son fundamentales para establecer tanto la función como el entorno en el que actúa.

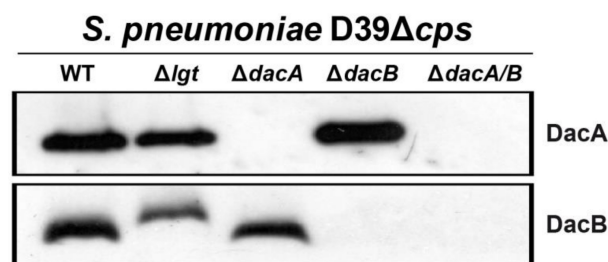
#### 5.5.4.1. Localización genómica de DacB

Análisis bioinformáticos revelaron que el gen *dacB* se encuentra en una región del genoma de *S. pneumoniae* altamente conservada, junto con genes de proteínas ribosomales (*rblK* y *rplA*) y una proteína de la familia HIT (que contienen una tríada de histidinas). No existe relación funcional entre estas proteínas y DacB, si bien es probable que *dacB* se exprese asociado a ellos en cadenas de mRNA policistrónico.

#### 5.5.4.2. Anclaje de DacA y DacB a la membrana y ensayos de expresión

La proteína DacA se encuentra anclada a la membrana lipídica mediante una hélice C-terminal anfipática (Morlot *et al.*, 2005), mientras que DacB es una lipoproteína, lo que requiere la existencia de una peptidasa de señal específica de lipoproteína (Lsp) y una diacilgliceril-transferasa (Lgt). Como se ha demostrado anteriormente, la ausencia de Lgt desemboca en cambios en la maduración y distribución de lipoproteínas a lo largo de la superficie de *S. pneumoniae* (Pribyl *et al.*, 2014). Así pues, con la intención de analizar la función tanto de DacA como de DacB, se han realizado experimentos de virulencia y reciclaje de peptidoglicano, así como mutantes deficientes en Lgt, en DacA, en DacB y en DacA y DacB al mismo tiempo. Además, estas pruebas se llevaron a cabo tanto en *S. pneumoniae* D39 como en su variante isogénica no encapsulada (D39 $\Delta$ *cps*).

La expresión de DacA y DacB se comporta como se espera en los mutantes deficientes en sus respectivos genes *dac*. En el mutante D39 $\Delta$ *cps* $\Delta$ *lgt* se expresan tanto DacA como DacB, pero en el caso de DacB se obtiene una banda de mayor peso molecular. Este comportamiento se correlaciona con el hecho de que, debido a la ausencia de Lgt, la proteína DacB aún contiene el péptido señal típico de lipoproteínas, que no ha sido eliminado por la Lgt, mientras que DacA, por no ser lipoproteína, mantiene su peso molecular (Figura 5.45).



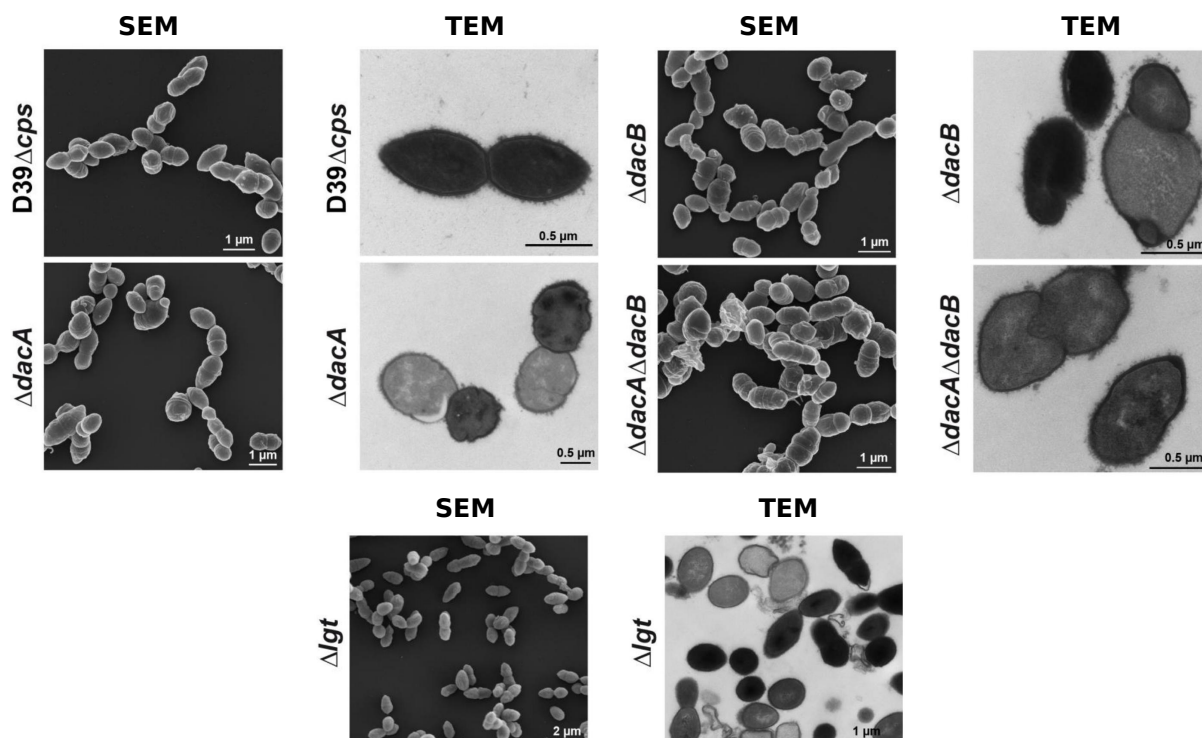
**Figura 5.45: Ensayos de expresión de DacA y DacB.** La proteína DacA se expresa en todas las cepas salvo en las que se ha delecionado *dacA*. DacB se expresa en todas las cepas salvo en las que se ha delecionado *DacB*. En el caso de la cepa D39 $\Delta$ *cps* $\Delta$ *lgt*, la proteína DacB presenta un tamaño ligeramente superior, debido a que realmente se trata de la pre-lipoproteína inmadura a la que no se le ha eliminado el péptido señal de anclaje a membrana.

#### 5.5.4.3. Caracterización fenotípica de células deficientes en DacA y DacB

Se realizaron observaciones mediante microscopía electrónica de transmisión y de barrido sobre células de D39 $\Delta$ *cps*, observándose la morfología habitual de *S. pneumoniae* (Figura 5.46). Los mutantes deficientes en *dacA* o *dacB* mostraron una morfología de longitud y ancho variable,



además de alteraciones en el proceso de división. En concreto, la falta de DacB ha generado células incapaces de separarse durante la división. Estos experimentos confirman la hipótesis de que DacB se localiza en el septo de división y participa en el proceso de división celular, propuesto por Barendt *et al.*, 2011. Los mutantes deficientes en *lgt* se mostraron ligeramente más alargados de lo normal.






















**Figura 5.46: Efecto de DacA, DacB y Lgt en la morfología celular.** Imágenes de SEM (Microscopía Electrónica de Barrido) y TEM (Microscopía Electrónica de Transmisión) de células de *S. pneumoniae* sometidas a delecciones de los diferentes genes de interés.

En lo que a susceptibilidad a antibióticos respecta, se determinaron las CMI para ampicilina, cefotaxima, ceftriaxona y vancomicina, para los diferentes mutantes diseñados. En general, la resistencia a  $\beta$ -lactamas no se vio afectada, pero los mutantes deficientes en alguna de las proteínas Dac o ambas mostraron mayor susceptibilidad a vancomicina, antibiótico glicopeptídico que interactúa con la fracción de D-alanil-D-alanina del peptidoglicano.

Por último, existía la posibilidad de que la cápsula polisacáridica se viera afectada por la delección de las proteínas Dac. Estudios mediante microscopía de fluorescencia y citometría de flujo demostraron la existencia de cantidades similares de cápsula tanto en *S. pneumoniae* D39 como en los fenotipos mutantes.

#### 5.5.4.4. Influencia en la composición del peptidoglicano

Para analizar la composición del peptidoglicano en los diferentes mutantes diseñados, así como en D39Δcps, se realizaron experimentos de digestión de la pared mediante la acción de LytA, combinados con cromatografías de exclusión molecular e identificación mediante espectrometría de masas (Figura 5.47).

Péptidos Estructura	Número	$\Delta cps$	$\Delta cps\Delta dacA$	$\Delta cps\Delta dacA$ $\Delta dacB$	$\Delta cps\Delta dacB$	$\Delta cps\Delta lgt$
	1	++	+			++
	2	+		+	++	
	3 (3*)	+	++	++		
	4/4'	+				+
	5/5'				+	
	6/6'				+	
	7/7'			+		
	8		+	+		+
	9	++	+			++
	10			+	++	
	11 (11*)		++	++		
	12A/B/A'/B'	++	+			++
	13/13'				+	
	14/14'				+	
	15/15'			+		
	16/16' (16*/16'*)		++	++		
	17/17'/17 <sup>ts</sup> /17''	+				+
	18/18'				+	
	19		+			

**Figura 5.47: Análisis de los fragmentos de peptidoglicano de *s. pneumoniae*.** Se muestran los diferentes péptidos obtenidos mediante digestión con LytA e identificados por espectrometría de masas. El número actúa como identificador. Los péptidos marcados con ++ han sido mayoritarios, mientras que los marcados con + son significativos. El color verde hace referencia a péptidos esperables del fenotipo salvaje, mientras que el color amarillo hace referencia a péptidos esperables en fenotipos deficientes en actividad D,D-carboxipeptidasa y el color rojo a fenotipos deficientes en actividad L,D-carboxipeptidasa.

Estos análisis confirmaron la estructura previamente conocida tanto para el fenotipo salvaje (Garcia-Bustos *et al.*, 1987; Severin & Tomasz, 1996; Bui *et al.*, 2012), como para los mutantes deficientes en los genes *dac* (Barendt *et al.*, 2011). De este modo se prueba que DacA se comporta como D,D-carboxipeptidasa y DacB como L,D-carboxipeptidasa.

En general, se observa que cada mutante produce los fragmentos peptídicos esperables en función de su fenotipo. En concreto, el mutante D39 $\Delta cps\Delta dacB$  produce fragmentos que

contienen el tetrapéptido, debido a su incapacidad de degradarlo a tripéptido. El componente mayoritario observado en los mutantes D39 $\Delta$ *cps* $\Delta$ *dacA* y D39 $\Delta$ *cps* $\Delta$ *dacA* $\Delta$ *dacB* contiene el pentapéptido, lo que coincide con la ausencia de actividad D,D-carboxipeptidasa y la necesidad de que DacA actúe antes que DacB.

No obstante, el mutante D39 $\Delta$ *cps* $\Delta$ *dacA* muestra algunos péptidos propios del fenotipo salvaje. Del mismo modo, el doble mutante D39 $\Delta$ *cps* $\Delta$ *dacA* $\Delta$ *dacB* también produce algunos péptidos típicos del fenotipo deficiente únicamente en DacB. Estas situaciones sugieren la posible existencia de una D,D-carboxipeptidasa adicional en *S. pneumoniae*, de menor afinidad que DacA, que se activa en ausencia de esta.

Otro caso especial es el del mutante D39 $\Delta$ *cps* $\Delta$ *lgt*, cuyo peptidoglicano presenta una composición habitual del fenotipo salvaje. Ya que la proteína Lgt está implicada en la maduración de DacB, estableciendo el enlace covalente entre la cisteína del motivo *LipoBox* y la membrana celular, es de esperar que un mutante deficiente en Lgt carezca de actividad L,D-carboxipeptidasa. Como se vio anteriormente, el mutante D39 $\Delta$ *cps* $\Delta$ *lgt* expresa la pre-prolipoproteína DacB inmadura (Figura 5.45). Este hecho, unido a la predicción de que el péptido señal actúa como hélice transmembrana, implica que la pre-prolipoproteína DacB podría translocarse y unirse a la superficie celular como una proteína de membrana estándar, mediante la hélice transmembrana, ejerciendo su actividad L,D-carboxipeptidasa con normalidad. No obstante, la cantidad de pre-prolipoproteína expresada por la cepa D39 $\Delta$ *cps* $\Delta$ *lgt* es significativamente menor que la expresada por la cepa D39 $\Delta$ *cps*, lo que explicaría el fenotipo ligeramente alargado de los mutantes deficientes en Lgt (Figura 5.46).

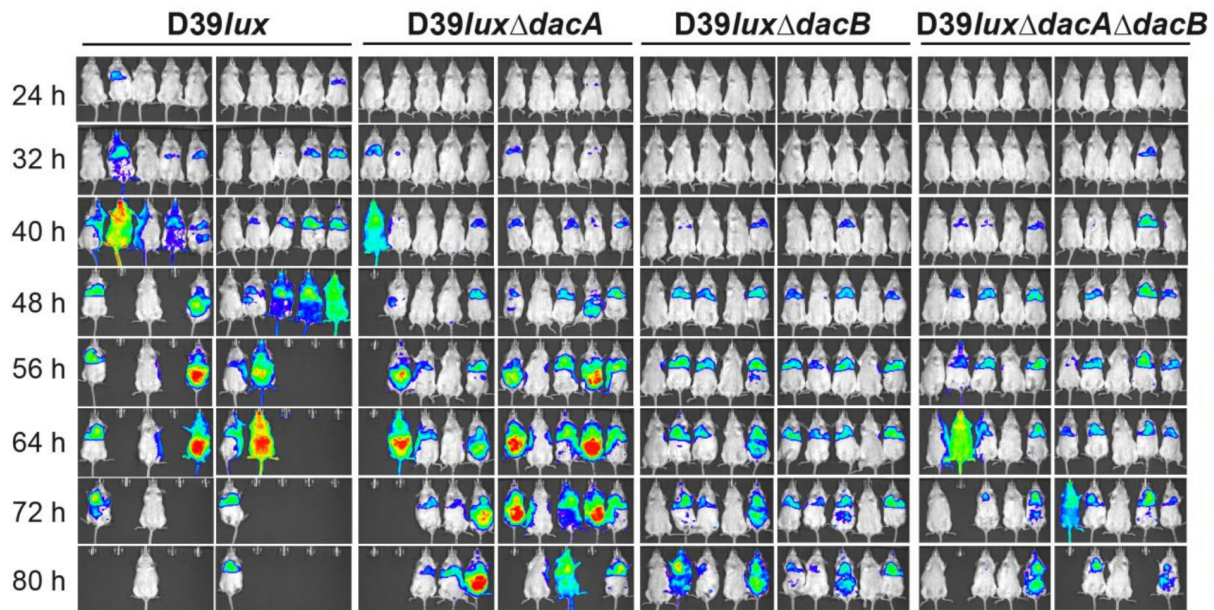
#### 5.5.4.5. Modelos de infección en ratones y supervivencia en macrófagos

Se realizaron ensayos de infección en ratones CD-1 con diferentes cepas de *S. pneumoniae* D39 bioluminiscente (D39*lux*), deficientes en *dacA*, en *dacB* y en ambos. La evolución de la infección se monitorizó a tiempo real. Mientras que los ratones inoculados con D39*lux* $\Delta$ *dacA* mostraban signos evidentes de infección a las 32 – 40 horas, los ratones a los que se les inoculó las cepas deficientes en DacB tardaron mucho más en mostrar los mismos efectos (Figura 5.48). En general, se observa que tanto DacA como DacB son fundamentales para mantener la virulencia de *S. pneumoniae*, si bien la pérdida de función por delección de DacB provoca un impacto mayor en la diseminación del patógeno que en los casos en los que se eliminó únicamente DacA.

Adicionalmente, se realizaron experimentos de supervivencia en macrófagos J774A.1, infectando tanto con D39 $\Delta$ *cps* como con las variantes isogénicas deficientes en DacA, en DacB y en ambas. Los resultados mostraron que la capacidad de supervivencia dentro de los macrófagos se reducía drásticamente en las cepas mutantes con respecto a D39 $\Delta$ *cps* (Figura 5.49).

#### 5.5.4.6. Adherencia

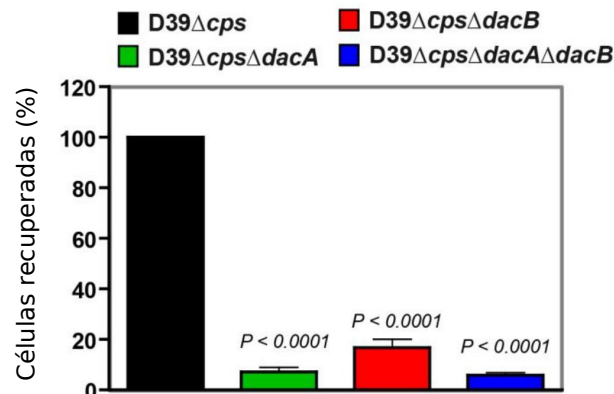
Los resultados de las infecciones *in vivo* demuestran una pérdida de la capacidad infectiva de *S. pneumoniae* al eliminar las proteínas DacA y DacB. Como se ha visto en las secciones anteriores, la pérdida de estas proteínas lleva a un cambio en la morfología del patógeno, pero



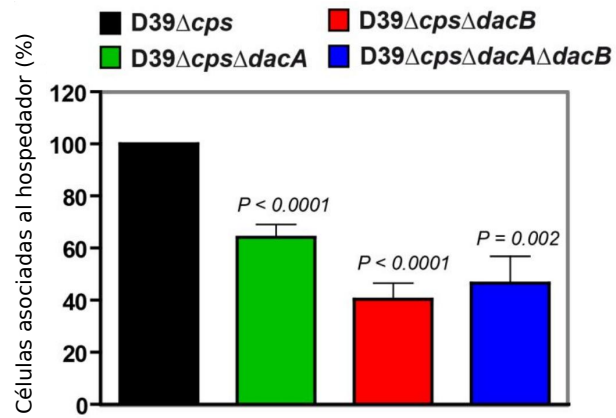
**Figura 5.48: Efecto de *DacA* y *DacB* en la infección en ratones.** Monitorización en tiempo real de la evolución de la infección por diferentes cepas de *S. pneumoniae* bioluminiscentes. La escala arcoiris de color indica la intensidad de la luminiscencia, recogida mediante un sistema IVIS Spectrum.

también podría verse reducida la capacidad de adhesión a las células del hospedador. De este modo, tanto la colonización como la posterior diseminación se dificultarían en gran medida.

Para analizar el efecto de las delecciones de *dacA* y *dacB* en el fenómeno de adhesión celular, se infectaron células epiteliales de alveolo pulmonar (A549) tanto con D39 $\Delta cps$  como con los mutantes isogénicos  $\Delta dacA$ ,  $\Delta dacB$  y  $\Delta dacA\Delta dacB$ . En general se observó una disminución de la adherencia en los mutantes respecto a D39 $\Delta cps$ , hecho que, unido a los defectos en la división celular, lleva nuevamente a una disminución en la capacidad infectiva general de las cepas mutantes (Figura 5.50).



**Figura 5.49: Efecto de *DacA* y *DacB* en la incorporación en macrófagos.** El porcentaje de células recuperadas del interior de los macrófagos tras una infección de 30 minutos es sensiblemente menor en el caso de las cepas deficientes en *DacA* y *DacB* que en la cepa salvaje no encapsulada.



**Figura 5.50: Efecto de DacA y DacB en la adherencia al hospedador.** El porcentaje de células adheridas a células epiteliales humanas se cuantificó a las 3 horas del momento de infección, observando un número significativamente menor en el caso de las cepas mutantes con respecto a la salvaje no encapsulada.



---

# DETERMINACIÓN ESTRUCTURAL Y FUNCIONAL DE CbPL

---

## Índice

---

6.1	Caracterización de CbpL . . . . .	<b>154</b>
6.2	Obtención de la proteína soluble . . . . .	<b>154</b>
6.2.1	Clonación y expresión . . . . .	154
6.2.2	Purificación . . . . .	155
6.3	Cristalización . . . . .	<b>157</b>
6.4	Determinación estructural . . . . .	<b>157</b>
6.4.1	Huella peptídica de los cristales . . . . .	158
6.4.2	Determinación estructural de CbpL <sub>CBD</sub> . . . . .	158
6.5	Variantes del protocolo de purificación inicial . . . . .	<b>160</b>
6.5.1	Proteólisis de CbpL en función de la temperatura . . . . .	161
6.6	Análisis estructural de CbpL <sub>CBD</sub> . . . . .	<b>162</b>
6.6.1	Estructura general . . . . .	162
6.6.2	Sitios de unión a colina . . . . .	163
6.6.3	Empaquetamiento cristalino . . . . .	164
6.6.4	Agitación térmica . . . . .	165
6.6.5	Otras moléculas y solvente . . . . .	165
6.7	Relación estructura-función . . . . .	<b>166</b>
6.7.1	Dominios Excalibur y Lipoprotein_Ltp . . . . .	168
6.7.2	Dominio CbpL <sub>CBD</sub> . . . . .	169
6.7.3	Localización y anclaje a la pared . . . . .	172
6.7.4	Estudios microbiológicos y fisiológicos relacionados con CbpL . . . . .	173

---

## 6.1. Caracterización de CbpL

El objetivo de los experimentos descritos en las siguientes secciones es la determinación estructural y funcional de la proteína CbpL de *S. pneumoniae* TIGR4.

La proteína CbpL consta de 332 residuos, con un peso molecular teórico de 37,6 kDa y un pI teórico de 6,2. Los primeros 26 aminoácidos pertenecen a un péptido líder implicado en el transporte y orientación de la proteína en su ubicación final. La proteína funcional se divide en 3 dominios. El primero es el dominio Excalibur (pfam05901), que abarca desde la His31 hasta el Glu67. En la Ala91 empieza el dominio de unión a colina (CbpL<sub>CBD</sub>), que se extiende hasta la Ala273. Finalmente, desde la Asn283 hasta el extremo C-terminal se encuentra el dominio Lipoprotein\_Ltp (pfam07553). Las regiones intermedias que conectan los dominios entre sí no parecen presentar estructura secundaria alguna, sino más bien actuar como zonas que aportan flexibilidad a la estructura general (Figura 6.1).

10	20	30	40	50	60
MNKR <sup>gr</sup> LFSKMS	LVTLPILALF	SQSVLA <sup>gr</sup> EE <sup>gr</sup> NI	HFSSCKEAWA	NGYS <sup>az</sup> DIHEGE	PGYS <sup>az</sup> AKLDRD
70	80	90	100	110	120
HDG <sup>az</sup> VACE <sup>az</sup> LKN	APKGAFKAKQ	STAIQINTSS	ATTSGWVKQD	GAWYYFDGNG	NLVKNAWQGS
130	140	150	160	170	180
YYLKADGKMA	QSEWIYDSSY	QAWYYLKSDG	SYAKNAWQGA	YYLKSN <sup>az</sup> GKMA	QGEWVYDSSY
190	200	210	220	230	240
QAWYYLKSDG	SYARNAWQGN	YYLKSDGKMA	KGEWVYDATY	QAWYYLTSDG	SYAYSTWQGN
250	260	270	280	290	300
YYLKSDGKMA	VNEWVDGGRY	YVGADGVWKE	VQASTASSN	DS <sup>az</sup> NSEYSAAL	GKAKSYNSLF
310	320	330			
HMSKKRMYRQ	LTSDFDKFSN	DAAQY <sup>az</sup> AIDHL	DD		

**Figura 6.1:** Secuencia de la proteína CbpL de *S. pneumoniae* TIGR4. El dominio Excalibur se ha coloreado en naranja, el de unión a colina en azul y el Lipoprotein\_Ltp en amarillo. El péptido señal se representa en gris. En negro se muestran las zonas de unión entre dominios.

## 6.2. Obtención de la proteína soluble

La obtención de la proteína CbpL soluble se realizó según se describe en las siguientes secciones. Ya que el péptido líder no está presente en la proteína madura funcional, no se incluyó en la construcción, que contiene desde el Glu27 hasta el Asp332, con un peso molecular teórico de 34,6 kDa y un pI teórico de 5,4.

### 6.2.1. Clonación y expresión

La secuencia de CbpL se clonó en el plásmido pET-28b(+) (Novagen) utilizando los sitios de restricción *Nco*I y *Bam*HI. La transformación se realizó sobre células de *E. coli* BL21(DE3) mediante choque térmico. La expresión se inició con la preparación de un precultivo de las



células transformadas en 20 mL de medio LB con 50  $\mu\text{g/mL}$  kanamicina. Este precultivo se dejó creciendo durante 16 horas a una temperatura de 37 °C.

Posteriormente se inocularon 10 mL del precultivo en matraces que contenían 1 L de LB autoclavado. A continuación se añadió 1 mL de kanamicina a una concentración de 50 mg/mL a cada matraz y se incubaron a 37 °C durante unas 3 ó 4 horas, hasta alcanzar una densidad óptica de 0,5. En ese momento se indujo la producción de la proteína mediante la adición de 1 mL de IPTG a una concentración 1 M. El cultivo inducido se mantuvo a una temperatura de 16 °C durante 18 horas.

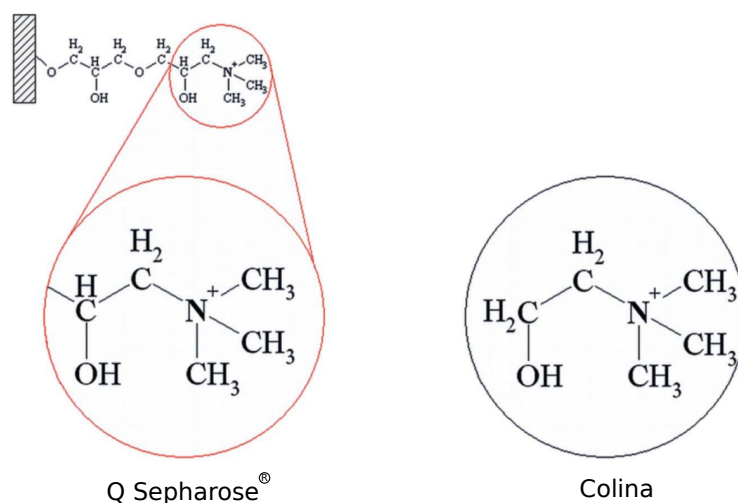
Tras esta última incubación, el cultivo se centrifugó a 4000 g durante 30 minutos para separar las células del sobrenadante, que fue eliminado. El precipitado obtenido de cada litro de cultivo se resuspendió en 10 mL de tampón fosfato sódico 20 mM pH 6,9 y se almacenó a -20 °C hasta su uso posterior.

### 6.2.2. Purificación

El primer paso para extraer la proteína expresada en la etapa anterior es producir la lisis de las células, ya sea por medios mecánicos o químicos. En este caso se optó por una lisis por sonicación por pulsos, evitando así un sobrecalentamiento del extracto. Tras la sonicación se procedió a una ultracentrifugación a 40000 g durante 50 minutos para retirar restos celulares y agregados. Posteriormente se añadió sulfato de estreptomicina para inhibir la síntesis de proteínas por los ribosomas que hayan quedado solubles y, tras 20 minutos, se realizó otra ultracentrifugación a 40000 g durante 50 minutos, buscando clarificar aún más la muestra.

El sobrenadante contiene tanto la proteína CbpL sobreexpresada como otras proteínas solubles que no han sido eliminadas en los pasos anteriores. Así pues, es necesario realizar un proceso de purificación más específico para retener la CbpL y eliminar la mayor cantidad de proteína contaminante posible. Para ello se realizó una cromatografía a baja presión en una columna con 5 mL de Q Sepharose® High performance (Sigma-Aldrich, EEUU). Los primeros protocolos de purificación de CBPs utilizaban columnas de DEAE celulosa por la afinidad de su matriz por los dominios de unión a colina (Sánchez-Puelles *et al.*, 1990). Posteriormente se mejoró el sistema de purificación al sustituir la DEAE celulosa por Q Sepharose® (Paterson *et al.*, 2006), mucho más parecida a las moléculas de colina que reconocen las CBPs y, por tanto, con mayor afinidad por los módulos de unión a colina (Figura 6.2). Así pues, aunque habitualmente estas columnas se utilizan para realizar cromatografías de intercambio iónico, en el caso de la CbpL se utilizó como columna de afinidad por el dominio de unión a colina. De esta forma se puede obtener la CbpL pura en un sólo paso cromatográfico.

La columna de Q Sepharose® se equilibró con 20 mM fosfato sódico pH 6,9, tampón en el que se encuentra la muestra y con el que se realizará la inyección y la entrada de la CbpL en la columna. A continuación, se pasó por la columna un tampón con 20 mM fosfato sódico pH 6,9 y 1,5 M NaCl, provocando la elución, por fuerza iónica, de todas aquellas proteínas que no son afines a la matriz. Posteriormente se rebajó el contenido en sal pasando un tampón que contiene 20 mM fosfato sódico pH 6,9 y 150 mM NaCl, hasta que deja de eluir proteína. En este punto,

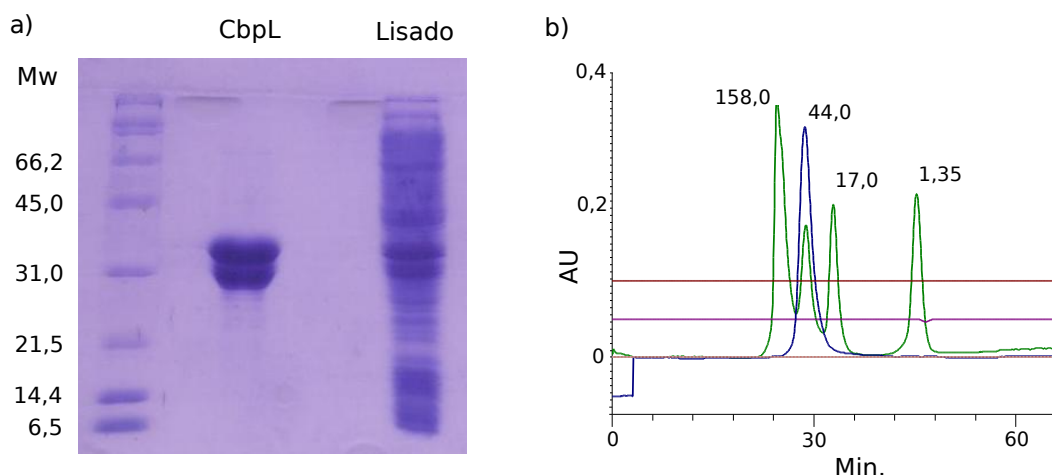


**Figura 6.2:** Similitud entre Q Sepharose<sup>®</sup> y la molécula de colina. Figura adaptada de Paterson *et al.*, 2006.

la proteína CbpL se encuentra retenida en la columna debido a la afinidad de sus dominios de unión a colina a la Q Sepharose<sup>®</sup>, mientras que el resto de contenido proteico de la muestra se ha eliminado. Para eluir la CbpL, se pasó un último tampón que contiene 20 mM fosfato sódico pH 6,9 y 0,5 M de cloruro de colina, creando un gradiente de colina en la columna. A medida que aumenta el gradiente, la proteína CbpL se libera progresivamente de la matriz, eluyendo de la columna y obteniéndose pura. Los tiempos de acción de cada tampón se determinaron de forma particular en cada caso, mediante la observación del cromatograma en tiempo real y realizando los ajustes necesarios.

Finalmente se realiza un último paso cromatográfico de exclusión molecular en columna HiLoad 16/600 Superdex<sup>®</sup> 200 PG preparativa en la que, además de aumentar el grado de pureza, se realizó un intercambio de buffer a 25 mM Tris-HCl pH 7,5, 150 mM NaCl y 140 mM colina, más adecuado que el fosfato sódico para la etapa de cristalización. La proteína eluida se concentró hasta valores óptimos para su cristalización mediante sucesivas centrifugaciones a 4000 g.

Como se observa en la Figura 6.3a, la purificación de CbpL genera dos bandas entre 30 y 35 kDa, que podrían corresponderse con dos proteínas diferentes, con una degradación parcial de la CbpL o con una migración diferencial en SDS-PAGE. Para determinar si la diferencia en tamaño de ambas bandas era significativa, se realizó una cromatografía de exclusión molecular en columna Superdex<sup>®</sup> 200 10/300 GL analítica (GE Healthcare, Suecia) que devolvió un único pico con un peso molecular en torno a 43 kDa (Figura 6.3b). Este experimento sugería que, probablemente, la doble banda observada en la SDS-PAGE se debiese a una agregación parcial de la muestra, por lo que se decidió continuar con los experimentos de cristalización. Por otro lado, la diferencia de 9 kDa observada entre el peso molecular teórico de la construcción de CbpL (34 kDa) y el peso molecular del pico obtenido en la exclusión molecular se puede deber a la morfología fibrilar de la proteína, mientras que los marcadores de peso molecular utilizados son globulares y su migración no es comparable a la de CbpL.



**Figura 6.3: SDS-PAGE y cromatografía de exclusión molecular de CbpL.** a) Gel SDS-PAGE mostrando el lisado celular y el eluido de la cromatografía en columna de Q Sepharose®. Se observa una banda doble en torno al marcador de 31,0 kDa. b) Cromatograma de la muestra de CbpL obtenida en la purificación, devolviendo un único pico simétrico y centrado en torno a 43 kDa (línea azul). Los marcadores de peso molecular (línea verde) utilizados son los Gel Filtration Standard de Bio-Rad (Hercules, CA, USA). Los valores mostrados sobre cada pico se corresponden al peso molecular en kDa.

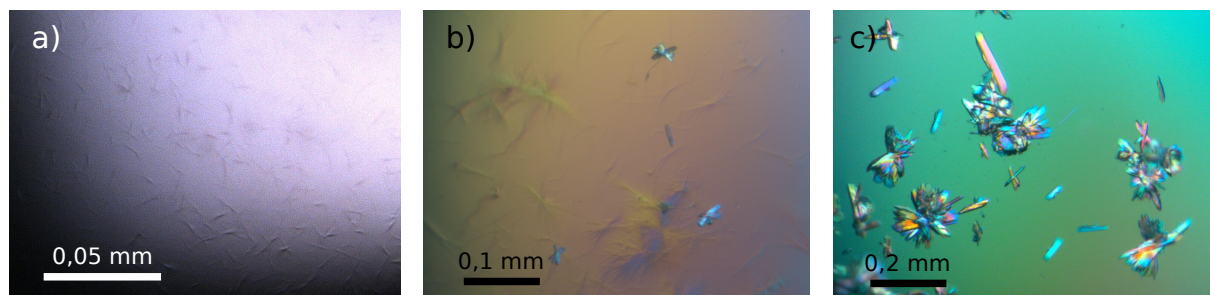
### 6.3. Cristalización

Los primeros ensayos de cristalización se realizaron con técnicas de alto rendimiento (Sección 3.2.5), utilizando los paquetes comerciales Crystal Screen, Crystal Screen 2, INDEX, JCSG, PACT Suite y JBScreen Classic 1-4 (Tabla 3.1). En cada ensayo de esta etapa se utilizaron 250 nL de CbpL a 40,8 mg/mL y 250 nL de cada condición de cristalización, a 18 °C, obteniéndose cristales extremadamente pequeños con forma de aguja a los 2 meses de iniciar los experimentos (Figura 6.4a). La condición de cristalización contenía 1,5 M sulfato amónico, 100 mM Tris-HCl pH 8,5 y 12 % (v/v) glicerol.

Debido a la presencia de sulfato amónico en la condición de cristalización, se realizó una optimización por difusión de vapor en gota colgante ensayando el paquete comercial JBScreen classic 6 (Tabla 3.1), cuyo componente principal es el sulfato amónico. Estos experimentos se realizaron añadiendo 1  $\mu$ L de CbpL a 40,8 mg/mL y 1  $\mu$ L de la condición de cristalización, a 18 °C. Se obtuvieron cristales con forma de barra y mayor tamaño bajo una condición que contenía 1,6 M sulfato amónico y 0,5 M LiCl (Figura 6.4b). La réplica manual del experimento anterior, con la finalidad de aumentar el número de cristales, sirvió indirectamente para obtener algunos cristales de mayor tamaño y mejor calidad, aptos para su resolución mediante cristalografía de rayos X (Figura 6.4c).

### 6.4. Determinación estructural

Los cristales obtenidos en la etapa anterior se crioprotegieron en 20 % (v/v) glicerol y sus espectros de difracción se recogieron en las líneas X06SA (PXI) y X06DA (PXIII) del SLS e ID29 del ESRF. La reducción de datos se realizó con los programas XDS, iMosflm y Aimless.



**Figura 6.4: Cristales obtenidos tras la purificación de CbpL.** a) Aguja pequeñas en una gota de 500 nL. b) Cristales de mayor tamaño, en una gota de 2  $\mu\text{L}$  de una condición de cristalización comercial. c) Cristales de mayor calidad bajo la misma condición del panel b, preparada de forma manual.

La celdilla unidad calculada durante el indexado ( $a = 30,85 \text{ \AA}$ ,  $b = 42,79 \text{ \AA}$ ,  $c = 70,62 \text{ \AA}$ ,  $\alpha = \gamma = 90^\circ$ ,  $\beta = 101,06^\circ$ ; Tabla 6.1) no permitía la existencia de la proteína CbpL en su interior, ya que su tamaño era inferior al mínimo necesario para que la CbpL formase un cristal. Así pues, el cristal obtenido no contenía la proteína CbpL y dos hipótesis eran posibles. La primera hipótesis supone la presencia en la muestra de una proteína diferente a la CbpL, contaminante de la purificación y de menor peso molecular, que desplazase a la CbpL y pudiera cristalizar en una celdilla unidad menor a la esperada. La segunda hipótesis implica una proteólisis parcial de CbpL, generándose fragmentos de menor tamaño que podrían cristalizar en celdillas unidad más pequeñas.

#### 6.4.1. Huella peptídica de los cristales

Con la intención de determinar qué hipótesis era correcta y poder afrontar la resolución estructural, se realizó una huella peptídica del contenido de los cristales que devolvería la secuencia de la proteína o fragmento cristalizado. Para ello se aislaron y trituraron varios cristales, disolviendo su contenido. A continuación se realizó una electroforesis en SDS, cortándose las bandas del gel para realizar una proteólisis *in situ* con tripsina. Los péptidos se extrajeron del gel y se concentraron para, posteriormente, someterlos a espectrometría de masas en el equipo 4800 MALDI-TOF/TOF (Applied Biosystems, EEUU) de la Unidad de Proteómica del Instituto de Investigación Sanitaria de Santiago de Compostela.

El resultado de este experimento indicó sin lugar a equívoco que la hipótesis correcta era la presencia de un fragmento parcial de CbpL (Figura 6.5). Así pues, tras la purificación de la proteína CbpL completa, se produce una proteólisis espontánea que fragmenta la proteína y elimina del experimento de cristalización los dominios Excalibur y Lipoprotein\_Ltp, quedando únicamente disponible el CbpL<sub>CBD</sub>.

#### 6.4.2. Determinación estructural de CbpL<sub>CBD</sub>

Así pues, se procedió a resolver los datos de difracción considerando que los cristales obtenidos hasta el momento contenían únicamente el CBD de CbpL. Las estadísticas de difracción del mejor cristal se muestran en la Tabla 6.1. El indexado, integrado y escalado de los diferentes cristales se realizó con los programas XDS, iMosflm, XSCALE y Aimless.

```

      10      20      30      40      50      60
MNKRLFSKMS LVTLPILALF SQSVLAEEENI HFSSCKEAWA NGYSIDIHEGE PGYSAKLDRD

      70      80      90      100     110     120
HDGVACELKN APKGAFAKAKQ STAIQINTSS ATTSQWVKQD GAWYYFDGNG NLVKNWQGS

     130     140     150     160     170     180
YYLKADGKMA QSEWIYDSSY QAWYYLKSDG SYAKNAWQGA YYLKSNGKMA QGEWVYDSSY

     190     200     210     220     230     240
QAWYYLKSDG SYARNAWQGN YYLKSDGKMA KGEWVYDATY QAWYYLTSDG SYAYSTWQGN

     250     260     270     280     290     300
YYLKSDGKMA VNEWVDGGRY YVGADGVWKE VQASTASSN DSNSEYSAAL GKAKSYNSLF

     310     320     330
HMSKKRMYRQ LTSDFDKFSN DAAQYAIHDL DD

```

**Figura 6.5: Huella peptídica sobre los cristales de CbpL.** La huella peptídica reveló que los cristales obtenidos tras la purificación de CbpL contenían aproximadamente la región de color rojo, correspondiente a la práctica totalidad del dominio de unión a colina de CbpL (zona roja y azul, CbpL<sub>CBD</sub>). En gris se indica el péptido líder, ausente en la construcción ensayada.

En base a la secuencia obtenida en la huella peptídica se estimó el tamaño del fragmento resultante en 21 kDa, lo que, unido a los datos de la celdilla unidad (Tabla 6.1), implica un coeficiente de Matthews de  $2,18 \text{ \AA}^3/\text{Da}$ , un porcentaje de solvente del 43,6 % y la presencia de una única molécula en la unidad asimétrica.

El reemplazo molecular se realizó con el programa Phaser, tomando como modelos una serie de fragmentos del dominio de unión a colina de la proteína CbpF (código PDB 2V04; Molina *et al.*, 2009). Tras la búsqueda de diferentes copias de cada modelo se encontraron un total de ocho sitios de unión a colina. Posteriormente, se mutaron con Chainsaw los residuos de cada sitio de unión a colina de CbpF por los correctos de CbpL y se acomodaron manualmente en la densidad electrónica, que predecía la mayoría de ellos con alta fidelidad.

**Tabla 6.1:** Datos y estadísticas de difracción del dominio de unión a colina de CbpL.

PARÁMETRO	CBPL <sub>CBD</sub>
Grupo espacial	P2 <sub>1</sub>
$a, b, c$ (Å)	30,85, 42,79, 70,62
$\alpha = \beta = \gamma$ (°)	90,0, 101,06, 90,0
Resolución (Å)	42,8 – 1,7 (1,73 – 1,7)
Longitud de onda (Å)	0,97857
Línea	PXI (SLS)*
Temperatura (K)	100
Nº de reflexiones	
Totales	130723
Únicas	19551
$\langle I/\sigma I \rangle$	27,5 (11,1)
$R_{\text{pim}}$	0,018 (0,056)
$CC_{1/2}$	0,99 (0,99)
Complejidad (%)	97,7 (96,8)
Multiplicidad	6,7 (7,0)

Los valores entre paréntesis hacen referencia a la última capa de resolución.

\* Se utilizó el detector Pilatus 6M.

El refinamiento del dominio de unión a colina de CbpL se realizó con phenix.refine, empezando por ciclos de cuerpo rígido y enfriamiento simulado para continuar posteriormente con sucesivos ciclos de reconstrucción manual en el programa Coot y refinamiento de coordenadas, factores térmicos y ocupaciones. En la Tabla 6.2 se muestran las estadísticas de refinamiento de esta estructura.

**Tabla 6.2:** Estadísticas de refinamiento de CbpL<sub>CBD</sub>.

PARÁMETRO	CBPL <sub>CBD</sub>
Resolución (Å)	36,4 – 1,7
$R_{\text{work}}/R_{\text{free}}$	0,15 / 0,19
Nº de átomos	
Proteína (no H)	1537
Colina	56
Glicerol	6
Sulfato	10
Agua	266
Factores B (Å <sup>2</sup> )	
Proteína	16,5
Colina	32,5
Glicerol	37,9
Sulfato	46,2
RMSD de enlaces	
Distancias (Å)	0,006
Ángulos (°)	0,886
Ramachandran (%)	
Favorables	99,5
Desfavorables	0,0
<b>Código PDB</b>	<b>4CNL</b>

Los valores de RMSD de enlaces y de Ramachandran son los esperables para unos datos de la calidad de los obtenidos con CbpL<sub>CBD</sub>. Los factores B medios de la proteína también se mantienen en los valores habituales para la resolución alcanzada, si bien el resto de moléculas distintas de la proteína presentan una agitación ligeramente superior.

## 6.5. Variantes del protocolo de purificación inicial

Siendo el objetivo principal la determinación estructural de la proteína CbpL y no únicamente su dominio de unión a colina, se realizaron diversos intentos de mejora del protocolo de purificación con la intención de evitar la proteólisis previa a la cristalización. La proteína extraída de cada nueva purificación se ensayaba con los paquetes comerciales de cristalización descritos en la sección 6.3, además del ya mencionado JBScreen Classic 6 y una serie de condiciones manuales dispuestas en torno a la condición de cristalización óptima de CbpL<sub>CBD</sub>. Además, se ensayaron diferentes concentraciones de proteína en cada caso.

En primer lugar, se añadieron inhibidores de proteasas tanto a los extractos celulares antes de proceder a su lisado como a los tampones utilizados durante todo el proceso de purificación. En concreto, se utilizaron dos inhibidores de proteasas, a saber, PMSF (200 µg/mL) y cOmplete

EDTA-free (en pastillas, Roche, Suiza), obteniendo mejores resultados en presencia del segundo compuesto. Se eligió el inhibidor libre de EDTA para no desestabilizar la unión a calcio del motivo Excalibur.

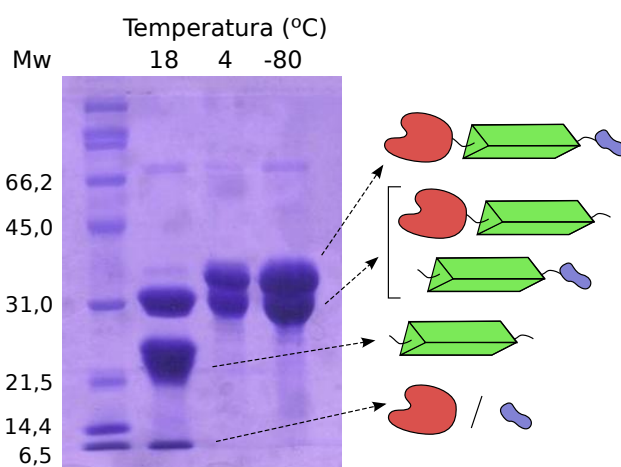
Si bien la purificación mejoraba notablemente tras la adición de los inhibidores de proteasas, se seguían obteniendo cristales de CbpL<sub>CBD</sub> 1 ó 2 meses después del inicio de los experimentos. Así pues, se realizaron modificaciones adicionales en el protocolo inicial de purificación que comprendían cambios de tampón (fosfato sódico por Tris-HCl o HEPES), metilación de la proteína o adición de DTT, con la finalidad de alterar el plegamiento y evitar la acción de proteasas. Además, debido a la posible unión de calcio, se realizaron algunas purificaciones añadiendo EDTA a los tampones y otras en las que se añadió un suplemento de CaCl<sub>2</sub>. En ninguno de los casos se obtuvieron cristales de la proteína CbpL completa.

Finalmente, se realizaron diversas purificaciones manteniendo una temperatura de 4 °C constante durante todo el proceso. Si bien la pureza y estabilidad de la muestra final eran las adecuadas para realizar los experimentos de cristalización (una única banda del tamaño esperable para la CbpL completa en SDS-PAGE), la cristalización a 4 °C no produjo ningún resultado que pudiera desembocar en posibles cristales.

### 6.5.1. Proteólisis de CbpL en función de la temperatura

Dado que la temperatura parecía ejercer un fuerte efecto en la proteólisis de CbpL, se realizó un experimento sencillo para determinar el comportamiento de esta bajo diferentes temperaturas habituales de trabajo. Una de las muestras de mayor calidad obtenidas en una purificación optimizada se alicuotó en tres fracciones y se dejó, durante 24 horas, a 18 °C, 4 °C y -80 °C respectivamente, analizándose posteriormente mediante SDS-PAGE (Figura 6.6).

El resultado muestra claramente que las muestras a 4 °C y -80 °C se han mantenido en las dos especies habituales de todas las purificaciones, con pesos de aproximadamente



**Figura 6.6: Efecto de la temperatura sobre la proteólisis de CbpL.** Gel SDS-PAGE con las muestras de CbpL tras 24 horas a una temperatura de 18, 4 y -80 °C. El diagrama muestra los posibles fragmentos en base al tamaño de cada banda, los pesos teóricos de cada dominio y la topología general de la proteína. El dominio Excalibur se representa en rojo, el dominio de unión a colina en verde y el dominio Lipoprotein\_Ltp en morado.



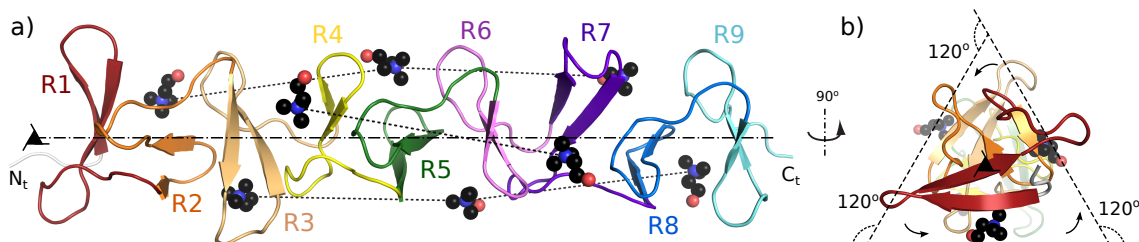
30 y 34 kDa respectivamente. Si se analiza el peso teórico de los diferentes dominios de CbpL, se puede extrapolar que la banda de 34 kDa se corresponde con la proteína CbpL entera, mientras que la banda cercana a 30 kDa haría referencia a la proteína sin el dominio Excalibur o sin el dominio Lipoprotein\_Ltp. En cambio, la muestra a 18 °C se ha degradado considerablemente, manteniendo la banda de 30 kDa y generando una nueva, mayoritaria, que abarca desde los aproximadamente 25 kDa hasta los 21 kDa, correspondiente al CbpL<sub>CBD</sub> aislado. Además, la última banda, de muy bajo peso molecular, podría corresponderse con el dominio Lipoprotein\_Ltp, con el Excalibur o con ambos.

## 6.6. Análisis estructural de CbpL<sub>CBD</sub>

La calidad de los mapas obtenidos permitió construir la estructura desde la Thr92 hasta la Ala273, lo que se corresponde con el dominio CbpL<sub>CBD</sub>. Además, como se describirá más adelante, se encontraron diversas moléculas de colina a lo largo de la estructura, unidas a los sitios específicos de unión a colina.

### 6.6.1. Estructura general

La estructura de CbpL<sub>CBD</sub> se asemeja a una superhélice  $\beta\beta - 3$  en solenoide compuesta por 9 repeticiones de una subestructura consistente en una horquilla  $\beta$  y una vuelta que enlaza con la siguiente repetición (Figura 6.7a). Las horquillas  $\beta$  se sitúan de forma perpendicular al eje de la superhélice, cuya longitud se sitúa en torno a los 85 Å. El apilamiento de las horquillas  $\beta$  a lo largo de la estructura se realiza en sentido antihorario (desplazándose del extremo N-terminal al C-terminal), con una rotación de 120° entre una horquilla y la siguiente (Figura 6.7b).



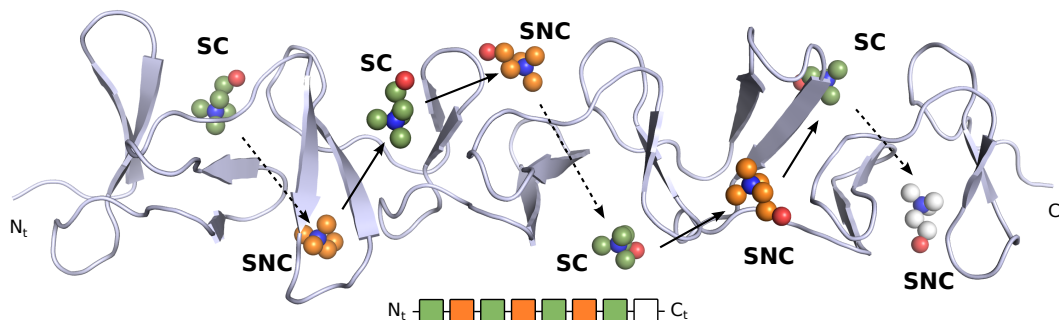
**Figura 6.7: Estructura general de CbpL<sub>CBD</sub>.** a) Estructura general de CbpL<sub>CBD</sub> coloreada por repeticiones (Rx) y mostrando las moléculas de colina unidas a los sitios específicos de unión a colina, siguiendo un eje de simetría  $3_1$ . Las líneas punteadas indican distancias de entre 26 y 28 Å. b) Vista a lo largo del eje longitudinal  $3_1$  en la que se aprecia la geometría triangular de CbpL<sub>CBD</sub> con ángulos de 120° entre cada repetición y su progreso en sentido antihorario. Las moléculas de colina se sitúan a lo largo de las tres caras del triángulo.

Cada repetición contiene una serie de residuos aromáticos que estabilizan internamente la superhélice mediante apilamiento de sus anillos. Además, otros residuos aromáticos quedan expuestos hacia el solvente de forma que dos repeticiones consecutivas generan un bolsillo hidrofóbico. Dicho bolsillo estabiliza las moléculas de colina mediante interacciones catión- $\pi$ , por lo que se denominan *sitios de unión a colina*. Estos sitios se sitúan alrededor del CbpL<sub>CBD</sub> siguiendo el eje longitudinal como si de un eje de simetría de orden  $3_1$  se tratara. De esta forma, los sitios de unión a colina se alinean a lo largo de 3 ejes paralelos al eje central de CbpL<sub>CBD</sub>, separándoles una distancia aproximada de entre 26 y 28 Å (Figura 6.7).



### 6.6.2. Sitios de unión a colina

Debido a la presencia de 9 repeticiones en la estructura y a que cada sitio de unión a colina está formado por una pareja de repeticiones consecutivas, se ha determinado la presencia de 8 sitios de unión a colina en CbpL<sub>CBD</sub>. De estos 8 sitios, 4 son denominados canónicos debido a su alta frecuencia de aparición en distintas proteínas de unión a colina, mientras que los otros 4 no canónicos no siguen las secuencias consenso observadas en los canónicos.



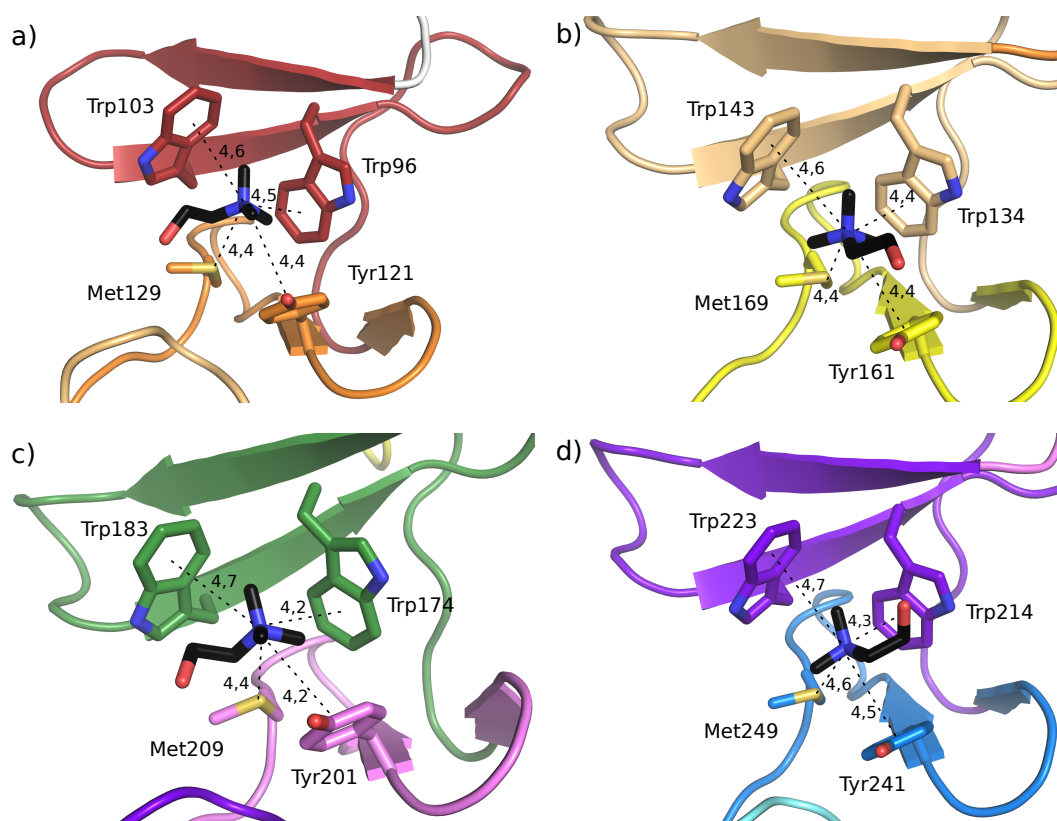
**Figura 6.8:** Distribución de sitios canónicos y no canónicos de unión a colina en CbpL<sub>CBD</sub>. Los distintos sitios de unión a colina se distribuyen de forma alterna a lo largo del dominio. Las flechas indican la consecución de sitios siguiendo la dirección  $N_t \rightarrow C_t$  (las flechas punteadas hacen referencia a giros por detrás de la estructura). Se puede apreciar el carácter helicoidal de la trayectoria formada por los diferentes sitios. Las moléculas de colina verdes hacen referencia a sitios canónicos (SC) mientras que las moléculas naranjas y blanca hacen referencia a dos tipos diferentes de sitios no canónicos (SNC).

#### 6.6.2.1. Sitios canónicos

Los sitios canónicos están formados por dos repeticiones consecutivas de unos 20 aminoácidos cada una. La secuencia consenso de cada repetición es  $GWXK-X_{4-5}-WYY-\phi-X_{3-5}GXM_{2-3}$ , donde X es cualquier residuo y  $\phi$  un aminoácido hidrofóbico. La horquilla  $\beta$  tiene una longitud de 12 a 14 residuos, seguida de una vuelta que contiene de 8 a 10 aminoácidos. La consecución de dos repeticiones da lugar al sitio de unión a colina, que consiste en un total de tres residuos aromáticos y una metionina. El primer residuo en participar en la interacción con la colina es un triptófano de la primera hebra  $\beta$  de la horquilla  $\beta$  de la primera repetición, seguido de otro triptófano de la segunda hebra  $\beta$  de la misma horquilla  $\beta$ . La segunda hebra  $\beta$  de la segunda repetición aporta una tirosina y, por último, se cierra el sitio con la metionina de la vuelta con la que termina la segunda repetición. En la Figura 6.9 se muestran los cuatro sitios canónicos de unión a colina presentes en CbpL<sub>CBD</sub>.

#### 6.6.2.2. Sitios no canónicos

Los sitios no canónicos de CbpL<sub>CBD</sub> están formados por dos repeticiones, al igual que los sitios canónicos. La diferencia recae en la composición de aminoácidos que interactúan con la colina, siendo cuatro los residuos aromáticos implicados en estos sitios. La primera repetición aporta un triptófano de la primera hebra  $\beta$  mientras que las tres tirosinas restantes provienen de la segunda repetición, concretamente del bucle de la horquilla  $\beta$ , de la segunda hebra  $\beta$  y de la vuelta final. En las Figuras 6.10a, 6.10b y 6.10c se muestran los tres sitios no canónicos de CbpL<sub>CBD</sub> que siguen esta configuración.



**Figura 6.9:** Sitios canónicos de unión a colina en CbpL<sub>CBD</sub>. Las repeticiones siguen el código de color de la Figura 6.7. Las moléculas de colina se han representado en varillas negras. Las líneas punteadas con valor numérico indican distancias medidas en Å. **a)** Primer sitio de unión a colina, formado por Trp96, Trp103, Tyr121 y Met129. **b)** Tercer sitio de unión a colina, formado por Trp134, Trp143, Tyr161 y Met169. **c)** Quinto sitio de unión a colina, formado por Trp174, Trp183, Tyr201 y Met209. **d)** Séptimo sitio de unión a colina, formado por Trp214, Trp223, Tyr241 y Met249.

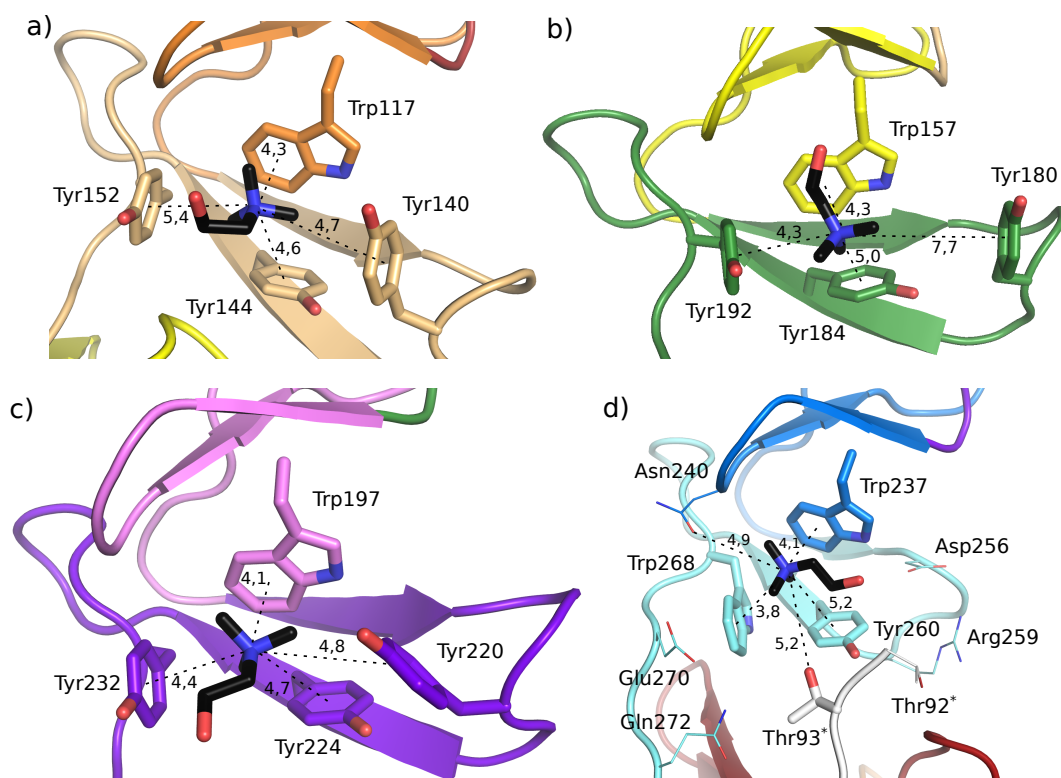
Además, se ha observado un último sitio de unión a colina que implica al Trp237 de la primera hebra  $\beta$  de la primera repetición, a la Tyr260 de la segunda hebra  $\beta$  de la segunda repetición y al Trp268 del extremo C-terminal de CbpL<sub>CBD</sub> cuya geometría se asemeja a la de un sitio canónico. La peculiaridad de este sitio surge al encontrar una treonina de una molécula vecina (Thr93) en el lugar que ocuparía la metionina canónica (Figura 6.10d).

### 6.6.3. Empaquetamiento cristalino

Como se indicó en la Sección 6.4.2, el cristal de CbpL<sub>CBD</sub> contiene una molécula por unidad asimétrica. Por tanto, ya que el grupo espacial es el monoclinico  $P2_1$ , hay un total de 2 moléculas por celdilla unidad.

Por convención, el eje binario helicoidal ( $2_1$ ) se corresponde con el eje  $b$  del sistema de coordenadas del cristal. Así pues, las vistas perpendiculares a los planos  $(b, c)$  y  $(a, b)$  permiten analizar el comportamiento de las moléculas a lo largo del eje  $2_1$ , observándose la rotación característica de  $180^\circ$  al alcanzar  $\frac{1}{2}b$  (Figuras 6.11a y 6.11b). La vista perpendicular al plano restante,  $(a, c)$ , permite apreciar que  $\beta \neq 90^\circ$  (Figura 6.11c).

Además de las generalidades ya descritas sobre la simetría de este cristal, el empaquetamiento de CbpL<sub>CBD</sub> presenta una peculiaridad. Como ya se indicó en la Sección 6.6.2.2, el último sitio



**Figura 6.10: Sitios no canónicos de unión a colina en CbpL<sub>CBD</sub>.** Las repeticiones siguen el código de color de la Figura 6.7. Las moléculas de colina se han representado en varillas negras. Las líneas punteadas con valor numérico indican distancias medidas en Å. **a)** Segundo sitio de unión a colina, formado por Trp117, Tyr140, Tyr144 y Tyr152. **b)** Cuarto sitio de unión a colina, formado por Trp157, Tyr180, Tyr184 y Tyr192. **c)** Sexto sitio de unión a colina, formado por Trp197, Tyr220, Tyr224 y Tyr232. **d)** Octavo sitio de unión a colina, formado por Trp237, Tyr260 y Trp268 principalmente. La Thr93 de la molécula vecina podría estar estabilizando directamente la colina, haciendo el papel de la metionina de los sitios canónicos. Otros residuos cercanos que podrían estabilizar la interacción se muestran en líneas. Los residuos de la molécula vecina se anotan con un asterisco.

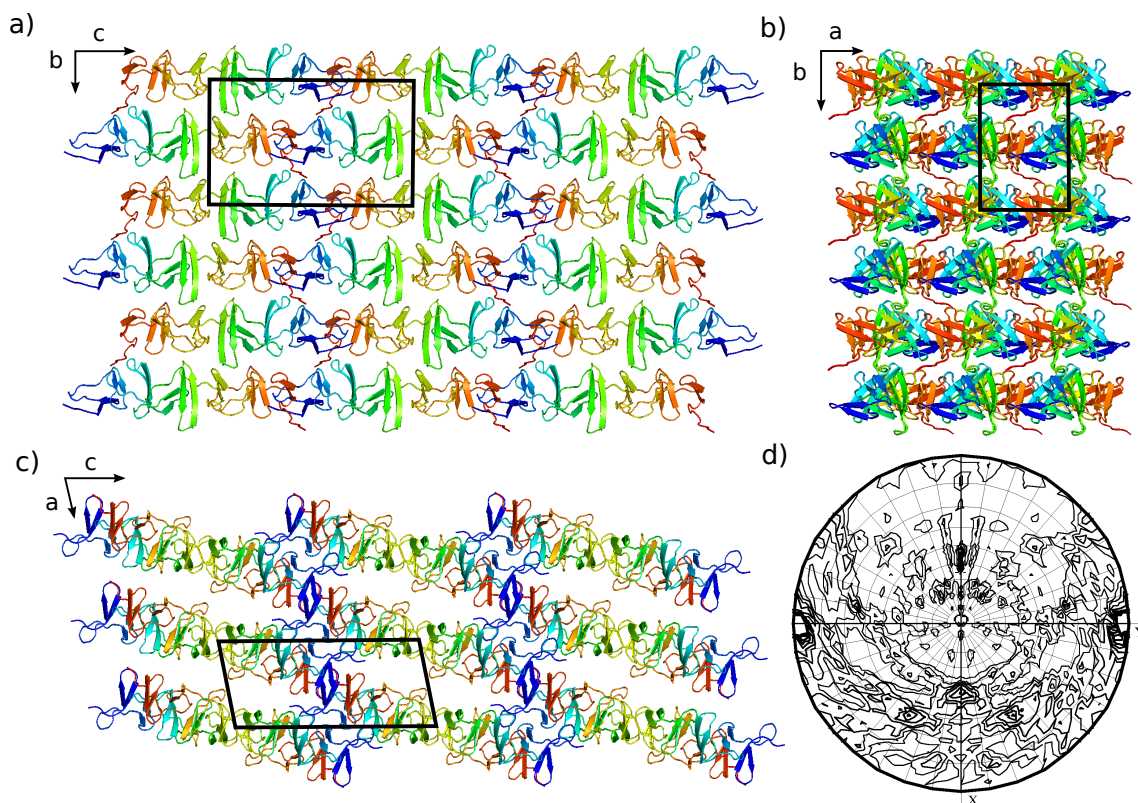
de unión a colina de CbpL<sub>CBD</sub> se compone de 3 residuos aromáticos con geometría similar a la de un sitio canónico pero en el que la metionina ha sido sustituida por una treonina de la molécula vecina. Esta molécula se encuentra perfectamente alineada con respecto a la original, de tal forma que el cristal se compone de cadenas alineadas de CbpL<sub>CBD</sub>, conectadas entre sí por el último sitio de unión a colina (Figura 6.12).

#### 6.6.4. Agitación térmica

Los factores B medios de la proteína se sitúan en 14,7 Å<sup>2</sup> para la cadena principal y en 18,2 Å<sup>2</sup> para las cadenas laterales, lo que implica cierta rigidez y elevada estabilidad y justifica el empaquetamiento observado y las estadísticas tanto de difracción como de refinamiento. Únicamente los bucles de algunas horquillas  $\beta$  muestran una agitación superior a la media, llegando a valores cercanos a 30 Å<sup>2</sup>, pero en ningún caso implican diferencias notables respecto a la flexibilidad general de la estructura (Figura 6.13).

#### 6.6.5. Otras moléculas y solvente

Como se ha visto anteriormente, en cada sitio de unión a colina se encontró una molécula de colina perfectamente estabilizada mediante interacciones catión- $\pi$  (Figura 6.14a). La presencia de



**Figura 6.11: Simetría y empaquetamiento de CbpL<sub>CBD</sub>.** La estructura se representa en código arcoiris para distinguir las moléculas individuales. El recuadro negro hace referencia a la celdilla unidad, indicándose además la orientación de cada eje en cada caso. **a)** Plano (*b*, *c*). **b)** Plano (*a*, *b*). **c)** Plano (*a*, *c*). **d)** Función de autorrotación calculada para  $\chi = 180^\circ$  mediante Molrep, mostrando los picos característicos de un grupo de Laue  $2/m$ .

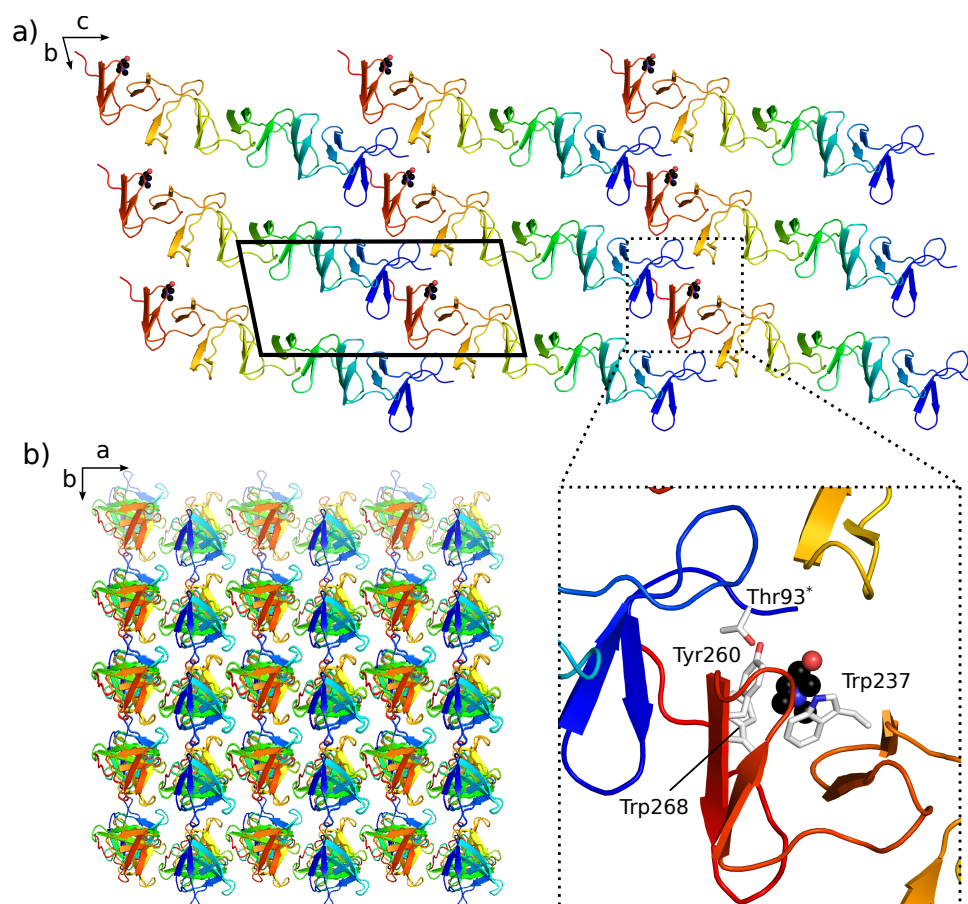
colina en el cristal se deriva de su uso para eluir la proteína durante el proceso de purificación. Sus factores B medios se sitúan en torno a  $32,5 \text{ \AA}^2$ , sin que existan diferencias significativas entre cada molécula (ya estén unidas a un sitio canónico o no canónico). No obstante, sí son significativas las diferencias en los factores B de cada átomo de cada colina ya que, mientras que los carbonos más cercanos a la proteína y el nitrógeno adquieren valores aproximados de  $20 \text{ \AA}^2$ , el oxígeno distal llega a superar los  $40 \text{ \AA}^2$  en la mayoría de los casos (Figura 6.14b). Estas discrepancias se justifican fácilmente por el hecho de que los átomos cercanos a la proteína son estabilizados mediante las interacciones catión- $\pi$  con el sitio de unión a colina mientras que el oxígeno se orienta hacia el solvente, sin presentar ninguna interacción con la proteína.

Además de las moléculas de colina, se modelaron dos sulfatos y un glicerol, con factores B medios de  $46,2$  y  $37,9 \text{ \AA}^2$  respectivamente. Finalmente, se añadieron 266 moléculas de agua.

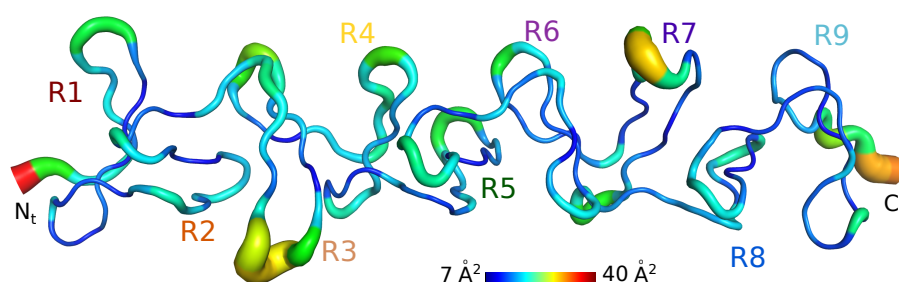
## 6.7. Relación estructura-función

Existe poca información previa disponible sobre la función de CbpL. Únicamente se ha observado su afinidad por colágeno, por elastina y por la proteína C reactiva (Frolet *et al.*, 2010), lo que sugiere una implicación a nivel de invasión y adhesión durante el proceso infeccioso de *S. pneumoniae*.



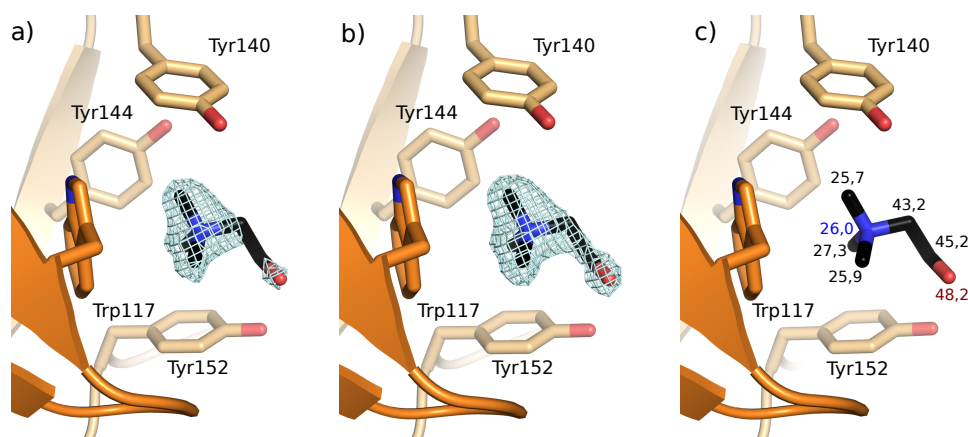


**Figura 6.12: Relación entre el empaquetamiento de CbpL<sub>CBD</sub> y el último sitio de unión a colina.** a) El cristal se compone de cadenas alineadas de CbpL<sub>CBD</sub>, conectadas entre sí por el último sitio de unión a colina. Mientras que una molécula aporta el extremo C-terminal y 3 residuos aromáticos (Trp237, Tyr260 y Trp268), la molécula vecina aporta el extremo N-terminal con la Thr93 (etiquetada como Thr93\*), quedando ambas regiones estables y en contacto. La colina que comparten se muestra en esferas negras. b) La existencia de contactos laterales complementa la estabilización de cada cadena. Además, la orientación de las cadenas de CbpL<sub>CBD</sub> se alterna por capas perpendiculares al eje *a*.



**Figura 6.13: Representación de factores de agitación térmica de CbpL<sub>CBD</sub>.** Los valores cercanos al azul reflejan baja agitación térmica, mientras que un virado hacia el rojo implica mayor agitación térmica y mayor flexibilidad.

Si bien en este trabajo se ha resuelto únicamente el CbpL<sub>CBD</sub>, es interesante analizar la relación entre la estructura y la función no sólo de este dominio, sino en el contexto de la proteína completa. Así pues, en las siguientes líneas se analizan tanto la estructura del dominio de unión a colina de CbpL en relación con otras CBPs, como las características generales de los dominios Excalibur y Lipoprotein\_Ltp. También se describirán los resultados funcionales

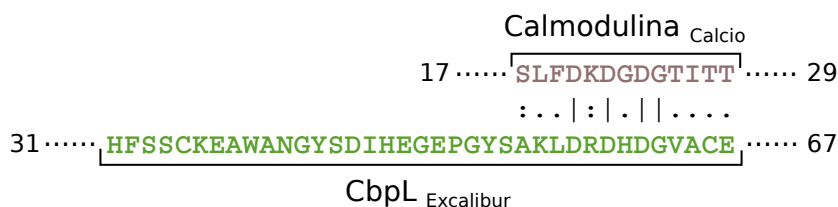


**Figura 6.14: Densidad electrónica y factores de agitación térmica de una molécula de colina.** a) Mapa de densidad electrónica  $2F_{\text{obs}} - F_{\text{cal}}$  a  $1,0 \sigma$ , correspondiente a la molécula de colina del segundo sitio de unión a colina de CbpL<sub>CBD</sub>. Se aprecia mayor densidad en la zona más estable de la molécula, mientras que el hidroxilo terminal aparece menos definido debido a su orientación hacia el solvente. b) Mapa de densidad electrónica de la misma colina del panel a, de tipo  $2F_{\text{obs}} - F_{\text{cal}}$ , calculado a  $0,8 \sigma$ . Pese a la mayor flexibilidad de la zona externa de la colina, los datos predicen la posición del hidroxilo con bastante fiabilidad. c) Valores atómicos de agitación térmica (factores B atómicos) de la misma colina del panel a, expresados en  $\text{\AA}^2$ , que corroboran la información aportada por el mapa de densidad electrónica.

realizados en este estudio y que pretenden proporcionar por primera vez una descripción del papel de esta proteína en la patogénesis de neumococo.

### 6.7.1. Dominios Excalibur y Lipoprotein\_Ltp

Como se ha comentado con anterioridad, la proteína CbpL contiene tres dominios bien diferenciados. En primer lugar, en el extremo N-terminal, se localiza el dominio Excalibur (residuos His31 a Glu67), que parece estar implicado en la unión extracelular de calcio (Rigden *et al.*, 2003). En general, los dominios Excalibur contienen el patrón DxDxDGxxCE, muy similar a la región de unión a calcio de la calmodulina (Figura 6.15). Así pues, se espera que los dominios Excalibur presenten un plegamiento de mano EF y sean capaces de unir  $\text{Ca}^{2+}$ . Este hecho concuerda con la interacción de CbpL con colágeno (Frolet *et al.*, 2010), ya que la unión a colágeno está mediada por motivos de unión a calcio (Wilson *et al.*, 2003).

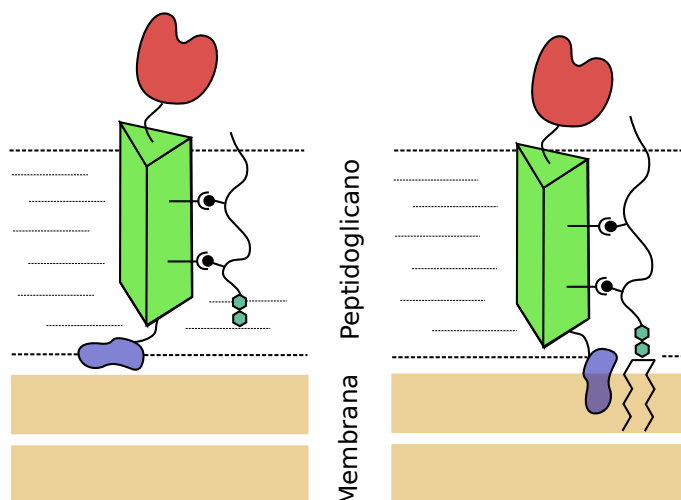


**Figura 6.15: Comparación entre el dominio Excalibur y el motivo de unión a calcio de calmodulina.** Se observa una alta similitud entre la secuencia consenso del dominio Excalibur de CbpL y el motivo de unión a calcio de calmodulina (código PDB 3CLN; Babu *et al.*, 1988).

Tras una región desordenada de 13 aminoácidos se sitúa el dominio de unión a colina (CbpL<sub>CBD</sub>), descrito anteriormente, que abarca desde la Ala91 hasta la Ala273. Finalmente, tras otra región desordenada de aproximadamente 9 residuos, se localiza el dominio Lipoprotein\_Ltp,

que abarca desde la Asn283 hasta el extremo C-terminal. Este dominio es característico de lipoproteínas de fagos encargadas de la transferencia del material genético hacia la célula hospedadora (Neve *et al.*, 1998). No obstante, la falta de literatura relacionada ha impedido obtener una descripción más precisa tanto de su estructura como de su función.

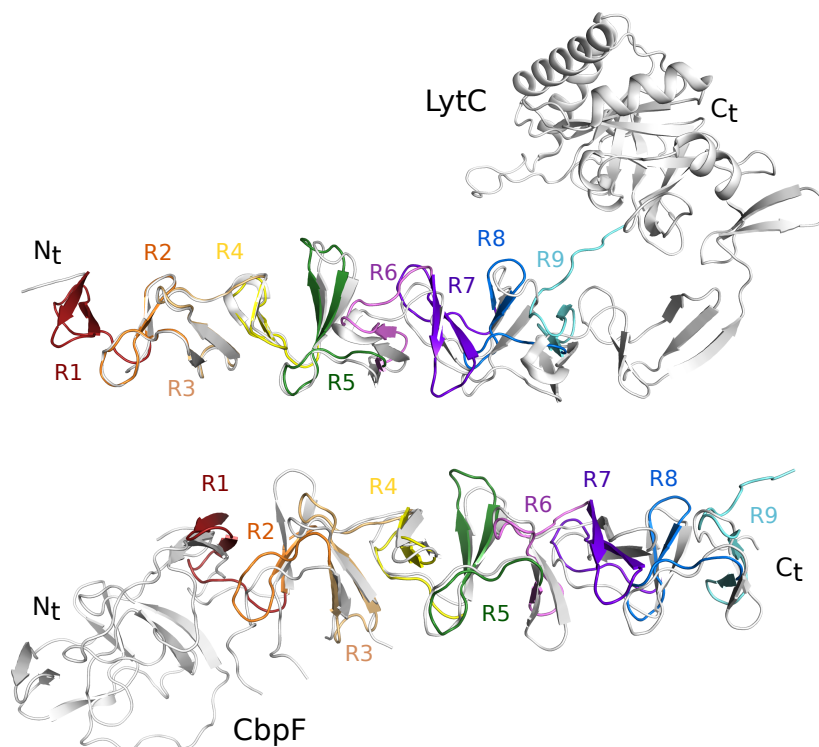
Es posible que el dominio Lipoprotein\_Ltp tenga como función interactuar con la membrana lipídica y orientar el resto de la proteína hacia el exterior, mientras que CbpL<sub>CBD</sub> ejerce un anclaje sólido al peptidoglicano. Bajo este contexto, CbpL se comportaría como lipoproteína sin el LipoBox habitual de las lipoproteínas clásicas (Oudega *et al.*, 1993). Otra posibilidad es que el dominio Lipoprotein\_Ltp interactúe con el fragmento lipídico de los ácidos lipoteicoicos, mientras que CbpL<sub>CBD</sub> se une a los fragmentos de colina. Además, la longitud de CbpL<sub>CBD</sub> es de aproximadamente 100 Å, lo que coincide con el grosor mínimo teórico aproximado del peptidoglicano de *S. pneumoniae*. De este modo, CbpL interactúa con la membrana lipídica mediante el Lipoprotein\_Ltp, se ancla al peptidoglicano mediante el CbpL<sub>CBD</sub> y expondría el Excalibur hacia la cápsula o medio extracelular (6.16).



**Figura 6.16: Orientación teórica de CbpL en la superficie celular.** El dominio Lipoprotein\_Ltp (morado) podría interactuar directamente con la membrana lipídica (izquierda) o bien con la zona de anclaje a membrana de los ácidos lipoteicoicos (derecha). En ambos casos, el dominio CbpL<sub>CBD</sub> (verde) se anclaría al peptidoglicano mediante interacción con los fragmentos de colina. El dominio Excalibur (rojo) se situaría hacia el exterior, sobresaliendo del peptidoglicano.

### 6.7.2. Dominio CbpL<sub>CBD</sub>

Como se indicó en la Sección 6.6.1, el dominio de unión a colina de CbpL consta de ocho sitios de unión a colina, cuatro canónicos y cuatro no canónicos. Estos sitios permiten el anclaje de CbpL a la pared bacteriana mediante enlaces de tipo catión- $\pi$  con los fragmentos de colina de los ácidos teicoicos y lipoteicoicos (Ma & Dougherty, 1997). El dominio de unión a colina análogo más cercano a CbpL<sub>CBD</sub>, calculado con el servidor Dali (Holm & Rosenström, 2010), es el de LytC (código PDB 2WWD, RMSD = 1,9 Å, identidad = 48 %; Pérez-Dorado *et al.*, 2010), seguido por CbpF (código PDB 2V05, RMSD = 3,1 Å, identidad = 37 %; Molina *et al.*, 2009) (Figura 6.17).



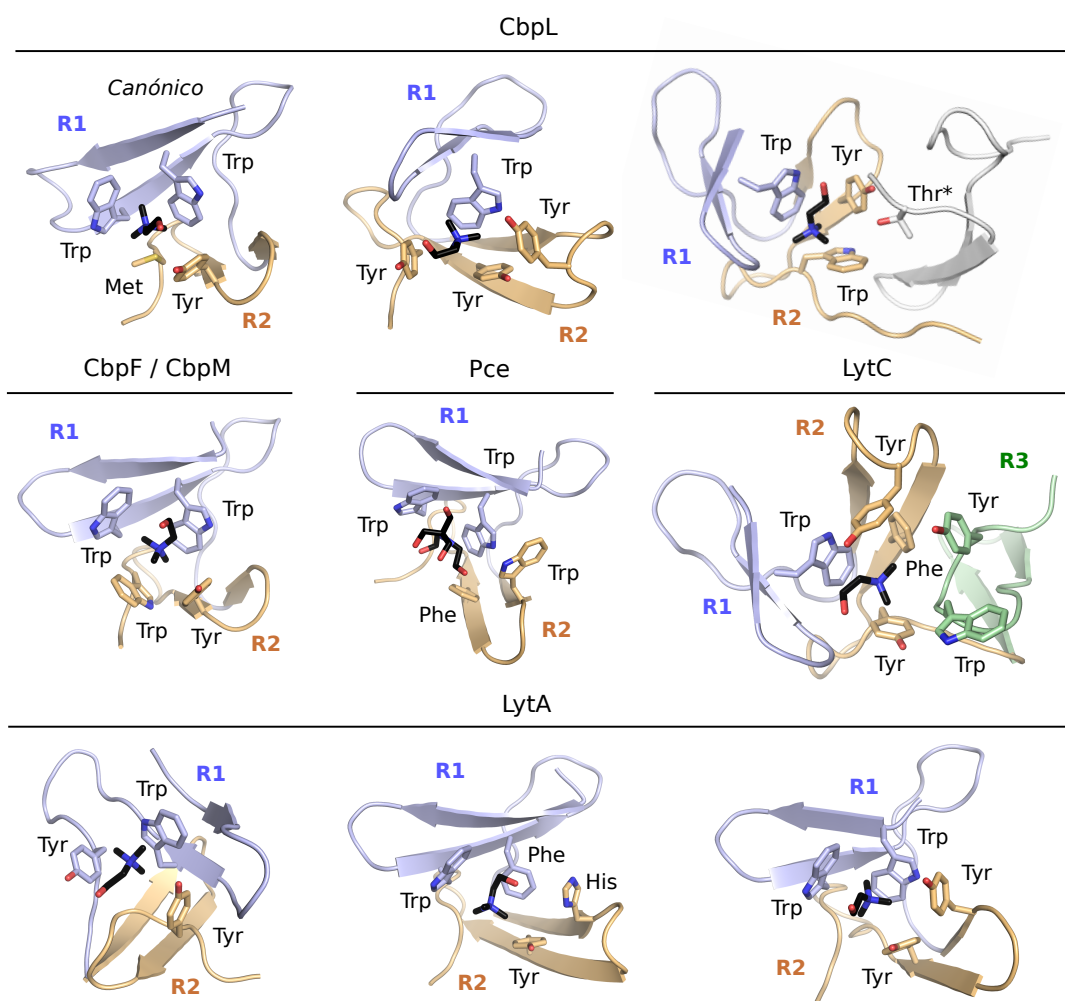
**Figura 6.17: Comparación estructural de CbpL<sub>CBD</sub> con LytC y CbpF.** Se observa que el plegamiento general del dominio de unión a colina se conserva tanto en LytC (código PDB 2WWD; Pérez-Dorado *et al.*, 2010) como en CbpF (código PDB 2V05; Molina *et al.*, 2009), ambas representadas en blanco. Se han etiquetado las repeticiones de CbpL<sub>CBD</sub> y los extremos N y C-terminales de LytC y CbpF.

De las 15 CBPs descritas, sólo se ha podido determinar la estructura tridimensional del dominio de unión a colina de 5 de ellas, a saber, LytA (códigos PDB 1HCX y 4IWT; Fernández-Tornero *et al.*, 2001; Mellroth *et al.*, 2014), Pce (código PDB 2BIB; Hermoso *et al.*, 2005), CbpM (código PDB 3HIA; sin publicación asociada) y las ya mencionadas LytC y CbpF.

En general, la composición de los sitios de unión a colina es muy variable. Su predicción en base a secuencia es compleja, ya que las repeticiones que los conforman constan de residuos aromáticos que no participan en la unión a colina y que pueden confundirse con los directamente implicados en la interacción. Así pues, la comparación de los sitios de unión a colina debe realizarse entre proteínas cuya estructura sea conocida. Como se vio en la Sección 6.6.2, los sitios observados con mayor frecuencia se denominan canónicos, presentan dos repeticiones con secuencia consenso  $\text{GWXK-X}_{4-5}\text{-WYY-}\phi\text{-X}_{3-5}\text{GXM}_{2-3}$  (siendo X cualquier aminoácido y  $\phi$  un residuo hidrofóbico) y constan de dos triptófanos aportados por la primera repetición estructural, seguidos de una tirosina y una metionina/leucina de la segunda repetición (Figura 6.18). Estos sitios canónicos aparecen tanto en CbpF (seis sitios), Pce (seis sitios), LytC (cinco sitios), LytA (dos sitios) y CbpM (dos sitios) como en CbpL (cuatro sitios).

Los sitios de unión a colina formados por repeticiones que no siguen la secuencia consenso se denominan sitios no canónicos. Estos sitios presentan una alta variabilidad entre las diferentes CBPs, pudiendo estar formados por dos o tres repeticiones que pueden implicar a tres, cuatro o seis residuos aromáticos en la unión a colina (Figura 6.18). Así pues, mientras que CbpF y





**Figura 6.18: Sitios de unión a colina.** Distribución de los diferentes tipos de sitios de unión a colina en las CBPs resueltas hasta la fecha. En CbpL, el residuo etiquetado con \* pertenece a una molécula vecina. En Pce, una molécula de Bis-Tris aparece unida al sitio, debido a su analogía estructural con la colina. LytC es la única proteína que presenta sitios formados por tres repeticiones.

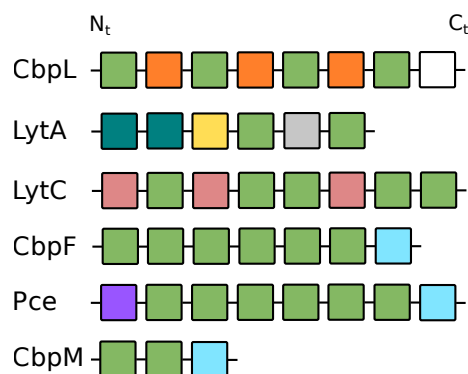
CbpM presentan un único sitio no canónico de unión a colina, formado por dos repeticiones que implican a dos triptófanos de la primera y a una tirosina y a otro triptófano de la segunda, Pce contiene además otro sitio no canónico formado por dos triptófanos de una repetición y un triptófano y una fenilalanina de la segunda. LytA contiene tres sitios no canónicos de unión a colina distintos, todos ellos formados por dos repeticiones. El primero está compuesto por un triptófano y una tirosina de la primera repetición y una tirosina de la segunda repetición, el segundo por un triptófano y una fenilalanina de la primera repetición y una histidina y una tirosina de la segunda, mientras que el tercero está formado por dos triptófanos de la primera repetición y dos tirosinas de la segunda. Por último, LytC presenta sitios no canónicos de unión a colina formados por tres repeticiones, implicando en la unión a un triptófano de la primera repetición, a dos tirosinas y una fenilalanina de la segunda repetición y a un triptófano y una tirosina de la tercera repetición. Esta última repetición contiene una secuencia característica formada por Gly–Tyr–Met–Ala (secuencia GYMA), también observada en LytB, que podría servir para reconocer los fragmentos de GlcNAc de los ácidos teicoicos y fortalecer el anclaje de la proteína a la pared (Pérez-Dorado *et al.*, 2010).

Como se ha descrito previamente, CbpL presenta cuatro sitios no canónicos exclusivos, de los cuales tres comparten configuración (Figura 6.18). El cuarto sitio no canónico se completa con la interacción con residuos de cadenas vecinas, por lo que no queda claro si nos encontramos frente a un nuevo tipo de sitio de unión a colina, también exclusivo de CbpL, o ante un artefacto cristalográfico debido al empaquetamiento.

Si bien la estructura general de los dominios de unión a colina sigue el plegamiento de una superhélice  $\beta\beta - 3$  en solenoide, la introducción de mutaciones en las repeticiones puede afectar al plegamiento y a la exposición de los sitios de unión a colina, alterando el anclaje de la proteína a la pared o modificando la función de las proteínas. Por ejemplo, la presencia de una hebra  $\beta$  adicional en una de las repeticiones de LytC genera un corte en la superhélice que desemboca en la estabilización y reorientación del dominio catalítico hacia el dominio de unión a colina (Figura 6.17). Esta conformación permite su activación, tras la interacción con CbpD, durante el proceso de fratricidio. Otro ejemplo de la influencia de la plasticidad de los sitios de unión a colina se encuentra en CbpF, en la que las inserciones en los sitios canónicos llevan a un cambio estructural drástico. En concreto, su extremo N-terminal se compone de seis repeticiones modificadas que generan un dominio capaz de inhibir tanto a LytC como a las lisozimas fágicas Cpl-1 y Cpl-7 (Figura 6.17) (Molina *et al.*, 2009).

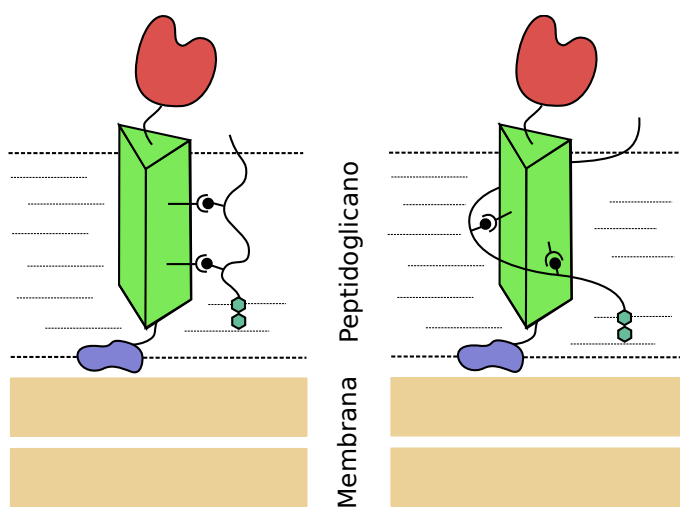
### 6.7.3. Localización y anclaje a la pared

Ya que todos los sitios de unión a colina, sean del tipo que sean, tienen por función anclarse al peptidoglicano mediante la interacción con la fosforilcolina de los ácidos teicoicos y lipoteicoicos, es de esperar que las diferencias entre ellos se relacionen con su afinidad de enlace. Además, la existencia de una distribución diferencial de sitios de unión a colina a lo largo de las diferentes proteínas de unión a colina descritas sugiere la existencia de patrones de sitios, que podrían actuar a modo de código, aportando a cada proteína una afinidad o ubicación determinada dentro de la pared de *S. pneumoniae* (Figura 6.19). Tal es el caso de LytB, que se localiza en los polos de la célula, o de CbpD, presente en el septo de separación celular. Otras, en cambio, se localizan uniformemente en toda la pared, como en el caso de CbpF.



**Figura 6.19: Distribución de sitios de unión a colina entre CBPs.** Se han coloreado en verde los sitios canónicos, presentes en todas las CBPs resueltas hasta la fecha. Los sitios no canónicos se han coloreado de manera diferencial dependiendo de su composición. Se puede observar que los sitios no canónicos son específicos de cada proteína salvo en el caso de CbpF, Pce y CbpM que comparten uno (formado por tres triptófanos y una tirosina).

Como se comentó con anterioridad, la presencia del dominio Lipoprotein\_Ltp sugiere que el dominio de unión a colina presenta el extremo C-terminal orientado hacia la membrana lipídica, quedando el N-terminal y el dominio Excalibur orientados hacia el exterior. Es importante destacar que CbpL es la única CBP que contiene el dominio Lipoprotein\_Ltp, por lo que es de esperar que su función aporte alguna característica diferencial respecto al resto de CBPs. En lo que respecta a la unión de CbpL<sub>CBD</sub> a los ácidos teicoicos y lipoteicoicos, caben dos hipótesis, a saber, que la distribución de las cadenas sea longitudinal, siguiendo alguna de las caras de CbpL<sub>CBD</sub>, o que se produzca una distribución en espiral, en la que los azúcares rodean al CbpL<sub>CBD</sub> y unen colinas de caras diferentes (Figura 6.20). Los análisis computacionales realizados en CbpF con el fin de determinar qué modelo era correcto (Molina, 2009) sugieren que la primera hipótesis es la adecuada. Ya que los dominios de unión a colina de CbpL y CbpF son altamente similares, es de esperar que el comportamiento de CbpL<sub>CBD</sub> también sea similar.



**Figura 6.20: Distribución teórica de las uniones a ácidos teicoicos y lipoteicoicos.** Los ácidos teicoicos y lipoteicoicos pueden seguir una distribución lineal a lo largo del CbpL<sub>CBD</sub> (izquierda) o rodear el dominio, implicando a otros sitios de unión a colina en el anclaje (derecha).

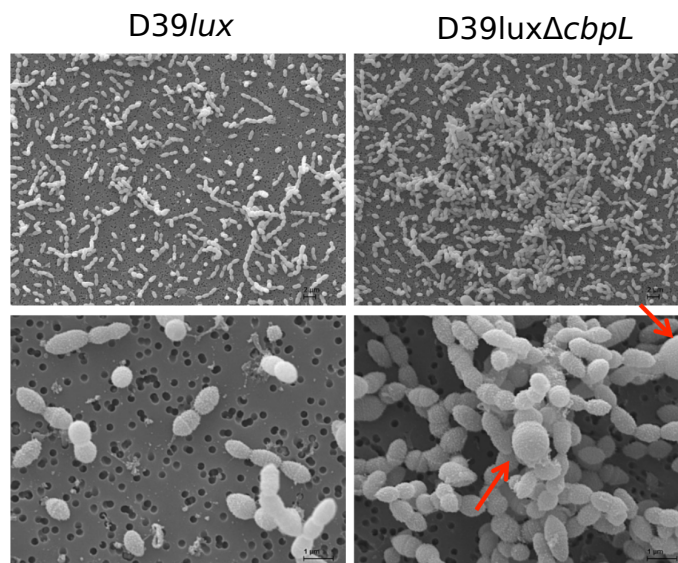
Finalmente, trabajos recientes han mostrado la importancia de la fosforilación de tirosinas expuestas en mecanismos de virulencia. En concreto, la fosforilación y defosforilación mediante la quinasa CpsD y la fosfatasa CpsB de una tirosina expuesta del dominio de unión a colina de LytA produce alteraciones en la capacidad de unir colina (Standish & Morona, 2013). La proteína CbpL presenta varias tirosinas expuestas en las cercanías de algunos sitios de unión a colina, por lo que no se descarta que sufra una serie de fosforilaciones y defosforilaciones que le permitan regular su anclaje a la pared, actuando en determinados momentos de la vida de la bacteria.

#### 6.7.4. Estudios microbiológicos y fisiológicos relacionados con CbpL

Los experimentos descritos a lo largo de esta sección han sido realizados por el grupo del Prof. Sven Hammerschmidt (Departamento de Genética Microbiana de la Universidad de Greifswald, Alemania) y se encuentran estrechamente relacionados con la estructura de CbpL<sub>CBD</sub> obtenida en este trabajo y son fundamentales para establecer tanto la función como el entorno en el que actúa.

#### 6.7.4.1. Caracterización fenotípica de células deficientes en CbpL

Se realizaron observaciones mediante microscopía electrónica de transmisión y de barrido sobre células de la cepa D39 bioluminiscente (*D39lux*) y sobre *D39luxΔcbpL*. Mientras que las células de *D39lux* mostraban la morfología habitual de *S. pneumoniae*, el mutante deficiente en CbpL mostraba alteraciones menores en el crecimiento de algunas unidades y una mayor agregación celular (Figura 6.21). No obstante, las diferencias en el crecimiento de ambas cepas no fueron lo suficientemente significativas como para asegurar que CbpL juegue un papel crucial en la remodelación del peptidoglicano o en el crecimiento celular.

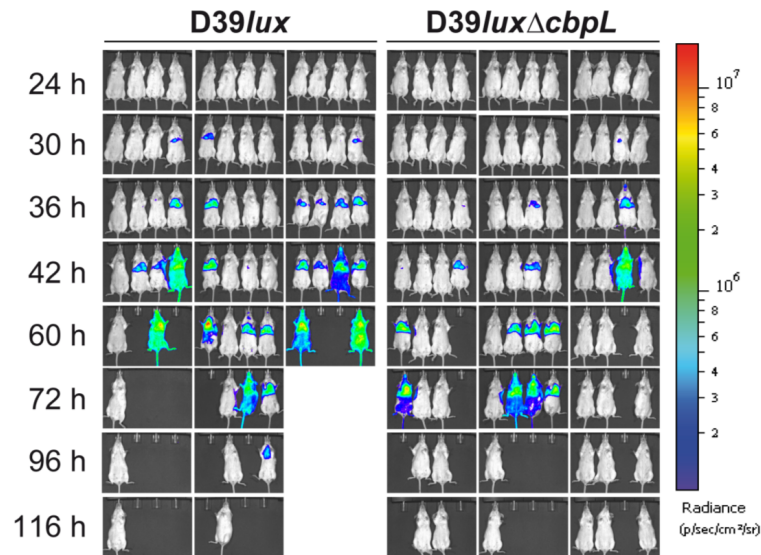


**Figura 6.21: Efecto de CbpL en la morfología celular.** Imágenes de SEM (Microscopía Electrónica de Barrido) y TEM (Microscopía Electrónica de Transmisión) de células de *S. pneumoniae* D39 bioluminiscentes (*D39lux*) y deficientes en CbpL (*D39luxΔcbpL*). Las flechas rojas señalan alteraciones puntuales en algunas células.

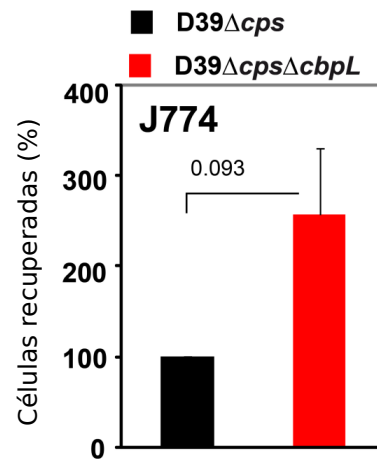
#### 6.7.4.2. Modelos de infección en ratones y supervivencia en macrófagos

Se realizaron ensayos de infección en ratones CD-1 con la cepa de *S. pneumoniae* D39 bioluminiscente (*D39lux*) y la deficiente en CbpL (*D39luxΔcbpL*). La evolución de la infección se monitorizó a tiempo real (Figura 6.22). Si bien todos los ratones empezaron a mostrar signos de infección a partir de las 30 horas, el proceso infectivo se retrasó ligeramente en los ratones inoculados con *D39luxΔcbpL*, cuya tasa de supervivencia es ligeramente superior a los inoculados con células *D39lux*.

Adicionalmente, se realizaron experimentos de supervivencia en macrófagos J774A.1, infectando tanto con la variante isogénica no encapsulada de *S. pneumoniae* D39 (*D39Δcps*) como con la variante deficiente en CbpL (*D39ΔcpsΔcbpL*). Sorprendentemente, se recuperaron más células en la infección con la cepa que no expresaba CbpL que en el experimento de control con *D39Δcps* (Figura 6.23). Este hecho sugiere que CbpL podría ser una diana crucial para el sistema inmune del hospedador y un punto débil del neumococo desde el punto de vista de su resistencia durante el proceso infectivo, ya que su ausencia favorece la supervivencia del patógeno e incluso su replicación.



**Figura 6.22: Efecto de CbpL en la infección en ratones.** Monitorización en tiempo real de la evolución de la infección por la cepa bioluminiscente de *S. pneumoniae* (D39 $lux$ ) y por la cepa deficiente en CbpL (D39 $lux\Delta cbpL$ ). La escala arcoiris de color indica la intensidad de la luminiscencia, recogida mediante un sistema IVIS Spectrum.



**Figura 6.23: Efecto de CbpL en la incorporación en macrófagos.** El porcentaje de células recuperadas del interior de los macrófagos tras una infección de 30 minutos es sensiblemente superior en el caso de las cepas deficientes en CbpL que en la cepa salvaje no encapsulada.



---

## CONSIDERACIONES FINALES

---

En este trabajo se han caracterizado estructural y funcionalmente tres proteínas implicadas en la biosíntesis y remodelado de la pared bacteriana de *S. pneumoniae*, así como en el proceso infectivo de este patógeno. A continuación se resumen los resultados más importantes de cada una de ellas y las implicaciones en la fisiología del neumoco.

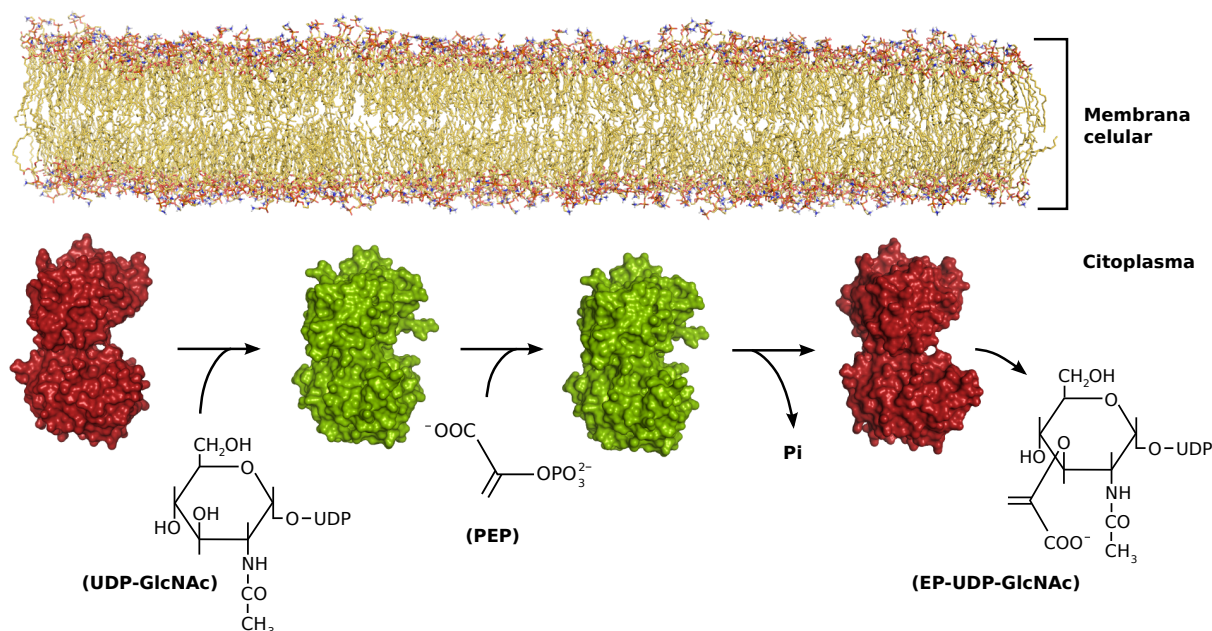
### 7.1. MurA y el inicio de la síntesis de peptidoglicano

La proteína MurA es la primera en actuar en la ruta de síntesis de peptidoglicano. En concreto, cataliza la conversión de UDP-GlcNAc en enolpiruvil-UDP-GlcNAc (EP-UDP-GlcNAc), que posteriormente será convertido en UDP-MurNAc por acción de la proteína MurB. En *S. pneumoniae* existen dos isoformas de la proteína MurA, a saber, MurA1 y MurA2. En este trabajo se ha cristalizado y determinado la estructura tridimensional de MurA1.

La estructura tridimensional de la enzima MurA1 consta de dos dominios globulares unidos por una bisagra doble, abarcando un total de 419 aminoácidos. La bisagra aporta flexibilidad a la estructura, que puede acercar o alejar sus dominios durante la catálisis. Así, inicialmente se encuentra en conformación abierta, exponiendo el sitio catalítico al solvente. Tras el reconocimiento y unión de la UDP-GlcNAc, primer sustrato de la reacción, sufre un cambio conformacional que provoca el acercamiento de los dominios, cerrando el sitio catalítico sobre la UDP-GlcNAc. Posteriormente se produce la entrada del fosfoenolpiruvato, del cual se incorporará, mediante la formación de un intermediario tetrahédrico, el fragmento enolpiruvil a la UDP-GlcNAc, liberando fosfato inorgánico. Una vez producida la reacción, la proteína recupera su conformación abierta, dejando el producto disponible para que la proteína MurB lo procese (Figura 7.1).

El mecanismo de catálisis implicaría a la Cys116, localizada en un lazo flexible que completa el sitio catalítico tras el cierre conformacional sobre la UDP-GlcNAc. No obstante, existen diferentes hipótesis acerca del mecanismo de reacción, no siendo siempre necesaria la intervención de dicha cisteína.





**Figura 7.1: Primera ruta de síntesis de peptidoglicano.** Inicialmente, la proteína MurA1 se encuentra en conformación abierta (rojo). El reconocimiento del primer sustrato (UDP-GlcNAc) induce el cierre de la conformación (verde). Posteriormente, el segundo sustrato (PEP) entra en el sitio catalítico y se produce la reacción, liberando fosfato inorgánico y el producto (EP-UDP-GlcNAc), lo que provoca nuevamente la apertura de la estructura.

Debido a que la proteína MurA1 es la encargada de iniciar la síntesis de peptidoglicano, ruta fundamental para la supervivencia de la bacteria, su inhibición o bloqueo podría ser letal. La fosfomicina, un antibiótico de amplio espectro, es capaz de unirse a la Cys116 e inhibir la actividad enzimática, bien por bloquear el residuo catalítico, bien por ocupar el sitio de unión a fosfoenolpiruvato. La mayoría de las cepas presentan heterorresistencia a fosfomicina, es decir, son capaces de producir subpoblaciones que crecen por encima de la CMI del antibiótico. No obstante, se ha observado que la cepa Hungary<sup>19A</sup> carece de heterorresistencia a fosfomicina. Entender las causas de este fenotipo es de gran importancia para desarrollar una terapia eficaz contra *S. pneumoniae*, dirigida al bloqueo de la síntesis de peptidoglicano, que evite la heterorresistencia a fosfomicina tan extendida entre las diferentes cepas de neumococo.

La única diferencia fenotípica observada entre las cepas heterorresistentes y la cepa Hungary<sup>19A</sup> es la presencia de la mutación Ala364Thr en la proteína MurA1 de *S. pneumoniae* Hungary<sup>19A</sup>. La determinación de la estructura tridimensional de esta enzima permitió analizar las diferencias estructurales existentes entre los dos fenotipos. La determinación estructural reveló que la mutación se localiza lejos del centro activo y su entorno se mantiene conservado en ambos fenotipos en lo que a orientación de las cadenas laterales respecta. Así pues, se descarta que el fenómeno de heterorresistencia a fosfomicina en *S. pneumoniae* se deba únicamente a esta diferencia estructural.

Un interesante resultado colateral de este trabajo se derivó de la cristalización de MurA1<sub>Hungary<sup>19A</sup></sub>. Esta enzima cristalizó en presencia de citrato, encontrándose esta molécula fuertemente unida al sitio de unión a fosfoenolpiruvato. Las interacciones que existen entre el citrato y el centro activo no sólo son superiores en número a las que existen con el



fosfoenolpiruvato, sino que además provocan el acercamiento de los dos dominios y, con ello, el cierre de la estructura. De esta forma la enzima quedaría bloqueada en su conformación cerrada e incapaz de incorporar la UDP-GlcNAc. Así pues, desde un punto de vista farmacológico, el citrato (o un análogo estructural) podría funcionar como nuevo antibiótico frente a la síntesis de peptidoglicano, lo que explicaría las recientes observaciones sobre las alteraciones letales que este compuesto produce en la pared del neumococo.

## 7.2. DacB, degradación de pared y división celular

La degradación de peptidoglicano es un proceso fundamental para el desarrollo de la bacteria. En *S. pneumoniae*, la división celular se realiza mediante la formación de un septo que divide a la célula por la mitad. Dicho septo se compone de una gruesa capa de peptidoglicano que debe sufrir una degradación parcial para liberar a las células hijas una vez hayan alcanzado el punto fisiológico adecuado. La degradación de la pared se utiliza también con la intención de liberar fragmentos de pared al medio para que actúen como factores de virulencia y para mantener el tráfico de determinados nutrientes hacia la célula.

El proceso de degradación de peptidoglicano en *S. pneumoniae* se inicia con la acción de la proteína DacA, una D,D-carboxipeptidasa que cataliza la hidrólisis del fragmento de D-Ala-D-Ala terminal de los péptidos de pared, dejando el tetrapéptido unido a las cadenas glicánicas y liberando unidades de D-Ala al medio. A continuación actúa DacB, una L,D-carboxipeptidasa que hidroliza el enlace entre la L-Lys y la D-Ala terminal del tetrapéptido, liberando la D-Ala al medio y dejando los tripéptidos unidos a las cadenas glicánicas (Figura 7.2).

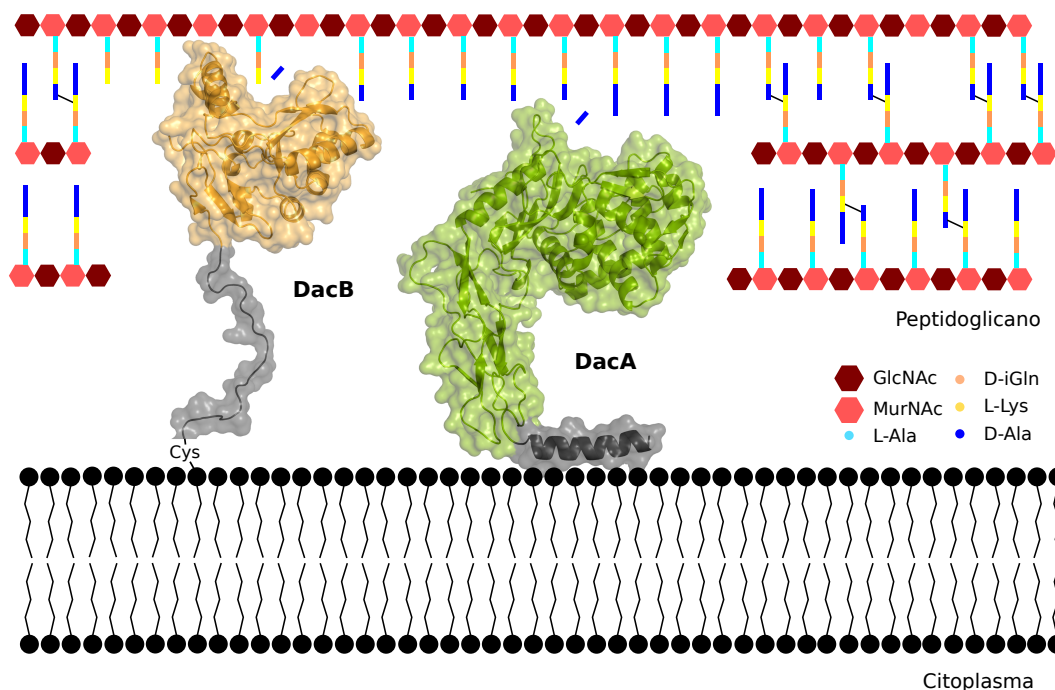
Los resultados obtenidos en este trabajo indican que DacB es una lipoproteína de superficie que contiene un péptido señal en el extremo N-terminal, seguido de una región flexible de unos 30 aminoácidos para terminar en un dominio catalítico de tipo globular, abarcando un total de 238 aminoácidos. El motivo *LipoBox* característico de las lipoproteínas se encuentra en el extremo C-terminal del péptido señal. Tras la eliminación del péptido señal, se establece un enlace covalente entre la cisteína del *LipoBox* y el diacilglicerol de los lípidos de la superficie de la membrana lipídica. De este modo se consigue un anclaje sólido a la membrana celular.

El centro catalítico se compone de un sitio de unión a zinc rodeado de elementos de estructura secundaria que conforman una cavidad capaz de albergar el tetrapéptido del peptidoglicano. Mientras que el plegamiento de la proteína se conserva en todas las estructuras determinadas, el lazo  $\beta_5 - \alpha_4$  cercano al sitio activo presentó dos conformaciones, abierta y cerrada, que regularían el reconocimiento del sustrato y el acceso a la cavidad catalítica. Así, la cavidad catalítica de DacB varía su tamaño en base a la orientación del lazo  $\beta_5 - \alpha_4$ , lo que permite que el dominio catalítico se acomode a los tetrapéptidos de la pared antes de realizar su degradación. Existen proteínas análogas en otros microorganismos (como VanXYc y VanXYg), con una maquinaria catalítica similar, que contienen un lazo de mayor longitud que cubre totalmente el sitio de unión a zinc. Estas proteínas presentan actividad D,D-carboxipeptidasa, en contraposición a DacB, que presenta actividad L,D-carboxipeptidasa. Nuestro análisis estructural ha mostrado

las bases moleculares de la especificidad de sustrato, que viene dada por las diferencias en el lazo  $\beta_5 - \alpha_4$  y en la composición de la cavidad catalítica.

La actividad carboxipeptidasa *in vivo* requiere de la actuación secuencial de la D,D-carboxipeptidasa DacA en primer lugar, seguida por la L,D-carboxipeptidasa DacB. Aunque ambas enzimas tienen una actividad catalítica similar, nuestro trabajo ha demostrado que su plegamiento estructural, maquinaria catalítica y regulación son radicalmente diferentes. Mientras que DacA presenta una maquinaria catalítica formada por tres motivos estructurales conservados entre los cuales se encuentra una serina catalítica, la proteína DacB utiliza un sitio de unión a zinc y un ácido glutámico para desestabilizar el enlace peptídico del sustrato y provocar su hidrólisis. Por otro lado, la rigidez de la estructura de DacA contrasta con la flexibilidad de DacB, que puede desplazar su dominio catalítico en un rango de hasta 100 Å sobre la superficie celular, mientras que DacA tiene limitado su rango de acción a una altura aproximada de 60 Å sobre la superficie celular. Las consecuencias de este mecanismo indirecto de regulación podrían relacionarse con la rotura del septo de división durante la separación de las células hijas, las cuales necesitan mantener un grosor mínimo de peptidoglicano.

Además, la presencia del lazo  $\beta_5 - \alpha_4$  en DacB y la posibilidad de dos conformaciones estructurales implica que la actividad L,D-carboxipeptidasa se encuentra regulada, mientras que la ausencia de un elemento similar en DacA sugiere que siempre está activa. Bajo este contexto, DacA eliminaría la D-Ala terminal de los pentapéptidos de forma sistemática. Posteriormente DacB degrada el tetrapéptido a tripéptido, de forma controlada. Aún se desconoce el mecanismo



**Figura 7.2: Modelo de degradación de peptidoglicano mediante DacA y DacB.** En primer lugar, DacA elimina las D-Ala terminales de los pentapéptidos. A continuación, DacB elimina la D-Ala restante, liberando tripéptidos. Se observa que la interacción de DacA con la membrana se realiza mediante una hélice anfipática, mientras que DacB, por ser lipoproteína, establece un enlace covalente entre la cisteína del motivo *LipoBox* y el fragmento de diacilglicerol de la membrana.

de regulación de DacB, así como si DacA es completamente activa o si también sufre algún tipo de regulación. Ya que ambas proteínas aparecen principalmente involucradas en la separación de las células hijas mediante la degradación del septo de división, es de esperar que alguna de las proteínas implicadas en la maquinaria de división celular ejerza algún tipo de control sobre ellas.

Ensayos realizados *in vivo* han demostrado que tanto DacA como DacB son esenciales para el correcto desarrollo del neumococo, ya que su ausencia genera fenotipos aberrantes, con alteraciones morfológicas graves. Además, los ensayos de infección en ratones han mostrado que estas proteínas son esenciales para la diseminación del patógeno y su virulencia. Por todo ello, el diseño de antibióticos contra estas carboxipeptidasas se presenta como una estrategia interesante en la lucha contra el neumococo.

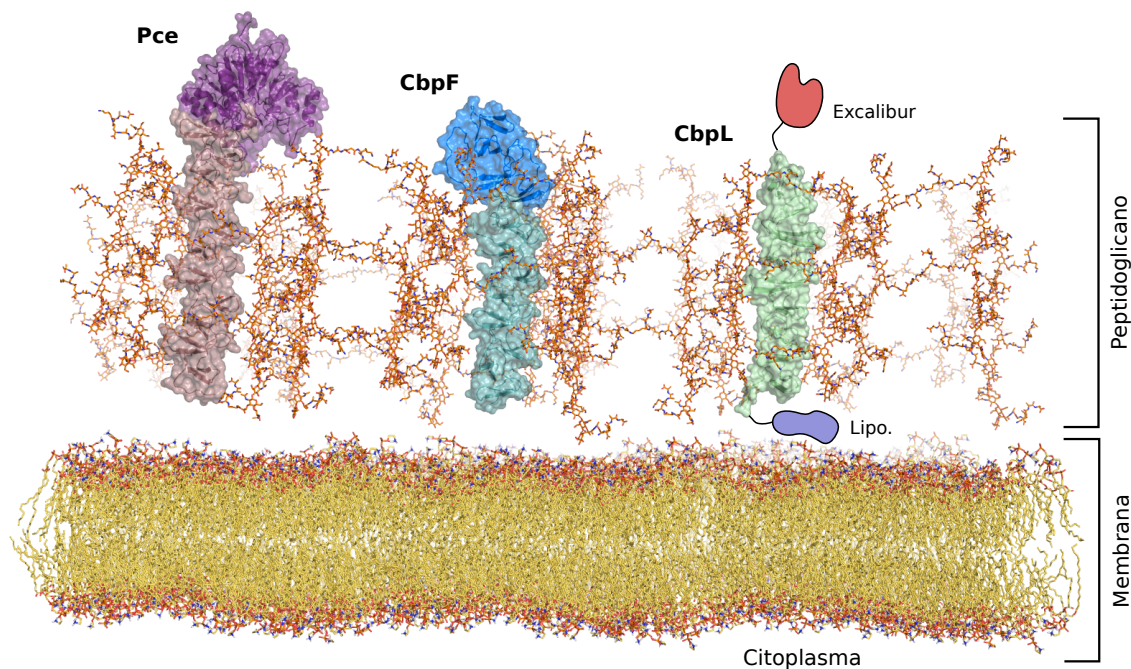
### 7.3. CbpL, anclaje a pared y adhesión al hospedador

Uno de los tipos más interesantes de proteínas de superficie de neumococo son las proteínas de unión a colina (CBPs), por ser exclusivas de *S. pneumoniae* y de fagos relacionados. Estas proteínas se caracterizan por su modularidad, conteniendo como mínimo un dominio de unión a colina, implicado en el anclaje a la pared bacteriana, y algún otro dominio con una función determinada. Si bien todas las CBPs tienen en común el dominio de unión a colina, las funciones que desempeñan a nivel fisiológico son muy diversas.

La proteína CbpL consta de tres dominios bien diferenciados. El primero, denominado Excalibur, de estructura desconocida, parece estar implicado en la interacción con calcio y la adhesión a elementos de la matriz extracelular de las células del hospedador. A continuación se localiza el dominio de unión a colina, que media el anclaje con la pared bacteriana. Por último, contiene un dominio denominado Lipoprotein\_Ltp, exclusivo dentro de las CBPs y de función desconocida, que podría estar implicado en la unión bien a la membrana citoplasmática, bien a la fracción lipídica de los ácidos lipoteicoicos. En total, 332 aminoácidos y un peso molecular de 37,6 kDa.

La determinación estructural del dominio de unión a colina de CbpL (CbpL<sub>CBD</sub>) reveló la presencia de ocho sitios de unión a colina, cuatro canónicos y cuatro no canónicos, alternos entre sí. Los sitios no canónicos observados son exclusivos de CbpL, no habiendo sido descritos en ninguna CBP con anterioridad.

Los sitios de unión a colina constan de una serie de residuos aromáticos que unen, mediante interacciones catión- $\pi$ , las fracciones de colina de los ácidos teicoicos y lipoteicoicos de la pared bacteriana. De esta forma se consigue que una proteína con una función determinada se ancle fuertemente a la superficie del microorganismo. En el caso de CbpL, probablemente implicada en el reconocimiento del hospedador y la adhesión, un anclaje sólido a la superficie celular es crucial para desempeñar su función. Como se ha dicho anteriormente, es probable que este anclaje se vea reforzado por la presencia exclusiva del dominio Lipoprotein\_Ltp, que podría interactuar con la membrana lipídica, complementando la unión de toda la estructura (Figura 7.3).



**Figura 7.3: Ubicación de las CBPs en la superficie celular.** El modelo de panel de abeja permite ubicar a las CBPs en los canales hexagonales que se forman en el peptidoglicano. Se observa que CbpL expone el dominio Excalibur de unión a calcio hacia el exterior, mientras que el dominio Lipoprotein\_Ltp (Lipo.) se encontraría enterrado entre el peptidoglicano y la membrana lipídica. Las proteínas Pce y CbpF se han añadido como referencia.

La distribución de los diferentes sitios de unión a colina a lo largo del dominio es un elemento diferencial de cada CBP y puede estar relacionado con su localización celular. Mientras que algunas CBPs se localizan de manera homogénea por toda la superficie del neumococo, otras CBPs se localizan exclusivamente en los polos celulares (como LytB) o en el septo de división (como CbpD). Además, existe una alta variabilidad en las funciones desempeñadas por estas proteínas, que están implicadas tanto en mecanismos de adhesión y reconocimiento del hospedador como en virulencia, regulación de otras proteínas de superficie o incluso autólisis. La disposición alternada de sitios canónicos y no canónicos en CbpL parece apuntar a una localización precisa en la pared.

Los ensayos de virulencia en ratones han mostrado que la ausencia de CbpL retrasa la diseminación de la infección y reduce ligeramente la mortalidad, lo que encaja con una función de adhesión y distribución del microorganismo hacia las zonas de infección. Además, la morfología celular apenas sufre cambios significativos, que pueden deberse más bien a una falta de contacto con la matriz extracelular que a alteraciones en la composición del peptidoglicano, como sucedía en el caso de DacA y DacB. Por último, la ausencia de CbpL dispara la supervivencia del patógeno en macrófagos, por lo que esta proteína podría jugar un papel fundamental en el reconocimiento por parte del sistema inmune y la eliminación de la infección.

---

---

## CONCLUSIONES

---

---



---

## CONCLUSIONES

---

A continuación se enumeran las principales conclusiones extraídas de este trabajo de investigación:

1. La proteína MurA1 de *S. pneumoniae* D39 se compone de dos dominios globulares unidos por una bisagra que aporta flexibilidad a la estructura, situándose el dominio catalítico entre ambas regiones. Un lazo flexible completa la maquinaria catalítica, participando en el reconocimiento de los sustratos y en la catálisis.
2. La proteína MurA1 de *S. pneumoniae* Hungary<sup>19A</sup> presenta el mismo plegamiento que MurA1<sub>D39</sub>, siendo el único elemento diferencial la mutación Ala364Thr en el dominio I.
3. La presencia de una molécula de citrato en el sitio de reconocimiento de fosfoenolpiruvato del centro activo de MurA1<sub>Hungary<sup>19A</sup></sub> provoca un bloqueo de la enzima en su conformación cerrada, actuando como inhibidor fuerte de su actividad.
4. El gen *murA1* de *S. pneumoniae* es el causante del fenotipo de heterorresistencia a fosfomicina común entre la mayoría de cepas de neumococo. El análisis estructural de MurA1<sub>D39</sub> y MurA1<sub>Hungary<sup>19A</sup></sub> (cepa no heterorresistente) demostró que dicho fenotipo no se debe a causa estructural directa.
5. La proteína DacB de *S. pneumoniae* TIGR4 es una lipoproteína de superficie con actividad L,D-carboxipeptidasa y formada por un péptido señal que contiene el *LipoBox*, seguido de una región flexible y del dominio catalítico.
6. El centro activo de DacB consta de un sitio de unión a zinc implicado en la catálisis y de un lazo cuya orientación define la forma de la cavidad catalítica y regula el reconocimiento del sustrato.
7. Los análisis estructurales de DacB permitieron identificar los residuos implicados en el reconocimiento específico del fragmento terminal L-Lys-D-Ala de los péptidos del peptidoglicano.
8. La proteína DacA elimina la última D-Ala del pentapéptido del peptidoglicano, devolviendo el tetrapéptido. La acción secuencial de DacB acorta este el tetrapéptido, liberando el cuarto residuo de D-Ala. Este mecanismo de acción, junto con la presencia de DacB en el septo de división celular, sugiere la existencia de una regulación de su actividad mediante elementos adicionales.

9. Las proteínas DacA y DacB participan activamente en la remodelación de pared y su delección resulta en alteraciones en la morfología de neumococo. Ambas enzimas presentan una maquinaria catalítica radicalmente diferente y sus estructuras tridimensionales apuntan hacia una funcionalidad *in vivo* distinta en cada caso.
10. Las proteínas DacA y DacB se comportan como elementos de virulencia fundamentales en el proceso de infección y adhesión a las células del hospedador, acelerando su invasión y muerte.
11. La proteína CbpL de *S. pneumoniae* TIGR4 es una proteína de unión a colina que consta de un dominio de unión a calcio, implicado en la adhesión a la matriz extracelular del hospedador, un dominio de unión a colina implicado en el anclaje a la pared celular y un dominio exclusivo entre las CBPs implicado en la interacción con los lípidos bien de la membrana celular, bien de los ácidos lipoteicoicos.
12. El dominio de unión a colina de CbpL (CbpL<sub>CBD</sub>) consta de ocho sitios de unión a colina, de los cuales cuatro son canónicos y cuatro no canónicos, alternos entre sí. Los sitios de unión a colina canónicos siguen la secuencia y composición consenso. Los sitios de unión a colina no canónicos son exclusivos de CbpL, no habiéndose observado en otras CBPs.
13. Diversos ensayos fisiológicos han mostrado la influencia de CbpL en la diseminación de la infección, hecho que se relaciona con su función de adhesión al hospedador. Además, CbpL funciona como elemento clave para la degradación del patógeno por los macrófagos del sistema inmune.



---

---

## BIBLIOGRAFÍA

---

---



---

## BIBLIOGRAFÍA

---

- Abendroth, J., Gardberg, A. S., Robinson, J. I., Christensen, J. S., Staker, B. L., Myler, P. J., Stewart, L. J. y Edwards, T. E. (2011) SAD phasing using iodide ions in a high-throughput structural genomics environment. *J Struct Funct Genomics*, **12**, 83–95.
- Abrahams, J. P. y Leslie, A. G. (1996) Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr D Biol Crystallogr*, **52**, 30–42.
- Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. y Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*, **66**, 213–221.
- Afonine, P. V., Grosse-Kunstleve, R. W., Urzhumtsev, A. y Adams, P. D. (2009) Automatic multiple-zone rigid-body refinement with a large convergence radius. *J Appl Crystallogr*, **42**, 607–615.
- Als-Nielsen, J. y McMorrow, D. (2011) *Elements of modern X-ray physics*. John Wiley & Sons.
- Anand, K., Pal, D. y Hilgenfeld, R. (2002) An overview on 2-methyl-2,4-pentanediol in crystallization and in crystals of biological macromolecules. *Acta Crystallogr D Biol Crystallogr*, **58**, 1722–1728.
- Anand, K., Schulte, A., Fujinaga, K., Scheffzek, K. y Geyer, M. (2007) Cyclin box structure of the P-TEFb subunit cyclin T1 derived from a fusion complex with EIAV tat. *J Mol Biol*, **370**, 826–836.
- Arndt, U. W., Crowther, R. A. y Mallett, J. F. W. (1968) A computer-linked cathode-ray tube microdensitometer for X-ray crystallography. *J Phys E Sci Instrum*, **1**, 510–516.
- Avery, S. V. (2006) Microbial cell individuality and the underlying sources of heterogeneity. *Nat Rev Microbiol*, **4**, 577–587.
- Babu, M. M., Priya, M. L., Selvan, A. T., Madera, M., Gough, J., Aravind, L. y Sankaran, K. (2006) A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J Bacteriol*, **188**, 2761–2773.
- Babu, Y. S., Bugg, C. E. y Cook, W. J. (1988) Structure of calmodulin refined at 2.2 Å resolution. *J Mol Biol*, **204**, 191–204.

- 
- Barends, T. R. M., de Jong, R. M., van Straaten, K. E., Thunnissen, A. M. W. H. y Dijkstra, B. W. (2005) *Escherichia coli* MltA: MAD phasing and refinement of a tetartohedrally twinned protein crystal structure. *Acta Crystallogr D Biol Crystallogr*, **61**, 613–621.
- Barendt, S. M., Sham, L.-T. y Winkler, M. E. (2011) Characterization of mutants deficient in the L,D-carboxypeptidase (DacB) and WalRK (VicRK) regulon, involved in peptidoglycan maturation of *Streptococcus pneumoniae* serotype 2 strain D39. *J Bacteriol*, **193**, 2290–2300.
- Bartual, S. G., Straume, D., Stamsås, G. A., Muñoz, I. G., Alfonso, C., Martínez-Ripoll, M., Håvarstein, L. S. y Hermoso, J. A. (2014) Structural basis of PcsB-mediated cell separation in *Streptococcus pneumoniae*. *Nat Commun*, **5**, 3842.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. y Leslie, A. G. W. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr*, **67**, 271–281.
- Bean, B. y Tomasz, A. (1977) Choline metabolism in pneumococci. *J Bacteriol*, **130**, 571–574.
- Beck, T., Krasauskas, A., Gruene, T. y Sheldrick, G. M. (2008) A magic triangle for experimental phasing of macromolecules. *Acta Crystallogr D Biol Crystallogr*, **64**, 1179–1182.
- Behr, T., Fischer, W., Peter-Katalinić, J. y Egge, H. (1992) The structure of pneumococcal lipoteichoic acid. Improved preparation, chemical and mass spectrometric studies. *Eur J Biochem*, **207**, 1063–1075.
- Bellamy, H. D., Snell, E. H., Lovelace, J., Pokross, M. y Borgstahl, G. E. (2000) The high-mosaicity illusion: revealing the true physical characteristics of macromolecular crystals. *Acta Crystallogr D Biol Crystallogr*, **56**, 986–995.
- Benvenuti, M. y Mangani, S. (2007) Crystallization of soluble proteins in vapor diffusion for X-ray crystallography. *Nat Protoc*, **2**, 1633–1651.
- Berejnov, V., Hussein, N. S., Alsaied, O. A. y Thorne, R. E. (2006) Effects of cryoprotectant concentration and cooling rate on vitrification of aqueous solutions. *J Appl Cryst*, **39**, 244–251.
- Bergfors, T. (2003) Seeds to crystals. *J Struct Biol*, **142**, 66–76.
- Bergfors, T. (2009) *Protein crystallization*. International University Line.
- Bergmann, S. y Hammerschmidt, S. (2006) Versatility of pneumococcal surface proteins. *Microbiology*, **152**, 295–303.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. y Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res*, **28**, 235–242.
- Berry, A. M., Lock, R. A., Hansman, D. y Paton, J. C. (1989) Contribution of autolysin to virulence of *Streptococcus pneumoniae*. *Infect Immun*, **57**, 2324–2330.
-

- 
- Beveridge, T. J.** (1981) Ultrastructure, chemistry, and function of the bacterial wall. *Int Rev Cytol*, **72**, 229–317.
- Bianco, C., Arena, F., Rossetti, B., Tordini, G., Migliorini, L., Galluzzi, P., Cerase, A., De Luca, A., Rossolini, G. M. y Montagnani, F.** (2014) First report of spondylodiscitis due to vancomycin heteroresistant *Staphylococcus capitis* in immunocompetent host. *J Infect Chemother*, **20**, 639–642.
- Boggon, T. J. y Shapiro, L.** (2000) Screening for phasing atoms in protein crystallography. *Structure*, **8**, R143–R149.
- Bourenkov, G. P. y Popov, A. N.** (2006) A quantitative approach to data-collection strategies. *Acta Crystallogr D Biol Crystallogr*, **62**, 58–64.
- Bracha, R., Chang, M., Fiedler, F. y Glaser, L.** (1978) Biosynthesis of teichoic acids. *Methods Enzymol*, **50**, 387–340.
- Bragg, W. L.** (1913) The diffraction of short electromagnetic waves by a crystal. *Proc Cambridge Philos Soc*, **17**, 43–57.
- Briles, E. B. y Tomasz, A.** (1970) Radioautographic evidence for equatorial wall growth in a gram-positive bacterium. Segregation of choline-3H-labeled teichoic acid. *J Cell Biol*, **47**, 786–790.
- Britton, D.** (1972) Estimation of twinning parameter for twins with exactly superimposed reciprocal lattices. *Acta Cryst*, **A28**, 296–297.
- Brown, E. D., Marquardt, J. L., Lee, J. P., Walsh, C. T. y Anderson, K. S.** (1994) Detection and characterization of a phospholactoyl-enzyme adduct in the reaction catalyzed by UDP-N-acetylglucosamine enolpyruvoyl transferase, MurZ. *Biochemistry*, **33**, 10638–10645.
- Brundish, D. E. y Baddiley, J.** (1968) Pneumococcal C-substance, a ribitol teichoic acid containing choline phosphate. *Biochem J*, **110**, 573–582.
- Brünger, A. T.** (1992) Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.
- Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K. y Jones, D. T.** (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res*, **41**, W349–W357.
- Bui, N. K., Eberhardt, A., Vollmer, D., Kern, T., Bougault, C., Tomasz, A., Simorre, J.-P. y Vollmer, W.** (2012) Isolation and analysis of cell wall components from *Streptococcus pneumoniae*. *Anal Biochem*, **421**, 657–666.
- Bunkóczi, G. y Read, R. J.** (2011) Improvement of molecular-replacement models with Sculptor. *Acta Crystallogr D Biol Crystallogr*, **67**, 303–312.
- Byczynski, B., Mizyed, S. y Berti, P. J.** (2003) Nonenzymatic breakdown of the tetrahedral (alpha-carboxyketal phosphate) intermediates of MurA and AroA, two carboxyvinyl transferases. Protonation of different functional groups controls the rate and fate of breakdown. *J Am Chem Soc*, **125**, 12541–12550.
-

- 
- Cartwright, K. A. V.** (2002) Epidemiology of meningococcal disease. *Hosp Med*, **63**, 264–267.
- Carugo, O.** (1999) Correlation between occupancy and B factor of water molecules in protein crystal structures. *Protein Eng*, **12**, 1021–1024.
- Cassidy, P. J. y Kahan, F. M.** (1973) A stable enzyme-phosphoenolpyruvate intermediate in the synthesis of uridine-5'-diphospho-N-acetyl-2-amino-2-deoxyglucose 3-O-enolpyruvyl ether. *Biochemistry*, **12**, 1364–1374.
- Catti, M. y Ferraris, G.** (1976) Twinning by merohedry and X-ray crystal structure determination. *Acta Crystallographica Section A*, **32**, 163–165.
- Chandra, N., Acharya, K. R. y Moody, P. C.** (1999) Analysis and characterization of data from twinned crystals. *Acta Crystallogr D Biol Crystallogr*, **55**, 1750–1758.
- Chayen, N. E.** (1997) The role of oil in macromolecular crystallization. *Structure*, **5**, 1269–1274.
- Chayen, N. E. y Saridakis, E.** (2008) Protein crystallization: from purified protein to diffraction-quality crystal. *Nat Methods*, **5**, 147–153.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. y Richardson, D. C.** (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*, **66**, 12–21.
- Chung, C. T., Niemela, S. L. y Miller, R. H.** (1989) One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc Natl Acad Sci U S A*, **86**, 2172–2175.
- Cianci, M., Helliwell, J. R. y Suzuki, A.** (2008) The interdependence of wavelength, redundancy and dose in sulfur SAD experiments. *Acta Crystallogr D Biol Crystallogr*, **64**, 1196–1209.
- Cowtan, K.** (2000) General quadratic functions in real and reciprocal space and their application to likelihood phasing. *Acta Crystallogr D Biol Crystallogr*, **56**, 1612–1621.
- Cowtan, K.** (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr*, **62**, 1002–1011.
- Cowtan, K.** (2010) Recent developments in classical density modification. *Acta Crystallogr D Biol Crystallogr*, **66**, 470–478.
- Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. y McPherson, A.** (1994) Screening and optimization strategies for macromolecular crystal growth. *Acta Crystallogr D Biol Crystallogr*, **50**, 414–423.
- Cundell, D. R., Gerard, N. P., Gerard, C., Idanpaan-Heikkila, I. y Tuomanen, E. I.** (1995) *Streptococcus pneumoniae* anchor to activated human cells by the receptor for platelet-activating factor. *Nature*, **377**, 435–438.
- Dauter, M. y Dauter, Z.** (2007) Phase determination using halide ions. *Methods Mol Biol*, **364**, 149–158.
-

- 
- Dauter, Z.** (1999) Data-collection strategies. *Acta Crystallographica Section D*, **D55**, 1703–1717.
- Dauter, Z.** (2003) Twinned crystals and anomalous phasing. *Acta Crystallogr D Biol Crystallogr*, **59**, 2004–2016.
- Dauter, Z.** (2005) Efficient use of synchrotron radiation for macromolecular diffraction data collection. *Prog Biophys Mol Biol*, **89**, 153–172.
- Dauter, Z., Dauter, M. y Dodson, E.** (2002) Jolly SAD. *Acta Crystallogr D Biol Crystallogr*, **58**, 494–506.
- de las Rivas, B., García, J. L., López, R. y García, P.** (2002) Purification and polar localization of pneumococcal LytB, a putative endo-beta-N-acetylglucosaminidase: the chain-dispersing murein hydrolase. *J Bacteriol*, **184**, 4988–5000.
- Delarue, M.** (2008) Dealing with structural variability in molecular replacement and crystallographic refinement through normal-mode analysis. *Acta Crystallogr D Biol Crystallogr*, **64**, 40–48.
- Deresinski, S.** (2013) The multiple paths to heteroresistance and intermediate resistance to vancomycin in *Staphylococcus aureus*. *J Infect Dis*, **208**, 7–9.
- Dessen, A., Mouz, N., Gordon, E., Hopkins, J. y Dideberg, O.** (2001) Crystal structure of PBP2x from a highly penicillin-resistant *Streptococcus pneumoniae* clinical isolate: a mosaic framework containing 83 mutations. *J Biol Chem*, **276**, 45106–45112.
- Diederichs, K. y Karplus, P. A.** (1997) Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol*, **4**, 269–275.
- Diederichs, K. y Karplus, P. A.** (2013) Better models by discarding data? *Acta Crystallogr D Biol Crystallogr*, **69**, 1215–1222.
- Dodson, E.** (1998) The role of validation in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*, **54**, 1109–1118.
- Dodson, E.** (2003) Is it jolly SAD? *Acta Crystallogr D Biol Crystallogr*, **59**, 1958–1965.
- Dramsi, S., Magnet, S., Davison, S. y Arthur, M.** (2008) Covalent attachment of proteins to peptidoglycan. *FEMS Microbiol Rev*, **32**, 307–320.
- Du, W., Brown, J. R., Sylvester, D. R., Huang, J., Chalker, A. F., So, C. Y., Holmes, D. J., Payne, D. J. y Wallis, N. G.** (2000) Two active forms of UDP-N-acetylglucosamine enolpyruvyl transferase in gram-positive bacteria. *J Bacteriol*, **182**, 4146–4152.
- Emsley, P., Lohkamp, B., Scott, W. G. y Cowtan, K.** (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*, **66**, 486–501.
- Engel, H., Mika, M., Denapaite, D., Hakenbeck, R., Mühlemann, K., Heller, M., Hathaway, L. J. y Hilty, M.** (2014) A low-affinity Penicillin-Binding Protein 2x variant is required for heteroresistance in *Streptococcus pneumoniae*. *Antimicrob Agents Chemother*, **58**, 3934–3941.
-

- 
- Eschenburg, S., Kabsch, W., Healy, M. L. y Schönbrunn, E.** (2003) A new view of the mechanisms of UDP-N-acetylglucosamine enolpyruvyl transferase (MurA) and 5-enolpyruvylshikimate-3-phosphate synthase (AroA) derived from X-ray structures of their tetrahedral reaction intermediate states. *J Biol Chem*, **278**, 49215–49222.
- Eschenburg, S., Priestman, M. y Schönbrunn, E.** (2005) Evidence that the fosfomicin target Cys115 in UDP-N-acetylglucosamine enolpyruvyl transferase (MurA) is essential for product release. *J Biol Chem*, **280**, 3757–3763.
- Eschenburg, S. y Schönbrunn, E.** (2000) Comparative X-ray analysis of the un-liganded fosfomicin-target MurA. *Proteins*, **40**, 290–298.
- Evans, G., Polentarutti, M., Carugo, K. D. y Bricogne, G.** (2003) SAD phasing with triiodide, softer X-rays and some help from radiation damage. *Acta Crystallogr D Biol Crystallogr*, **59**, 1429–1434.
- Evans, P.** (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr*, **62**, 72–82.
- Evans, P. R.** (1999) Some notes on choices in data collection. *Acta Crystallogr D Biol Crystallogr*, **55**, 1771–1772.
- Evans, P. R.** (2007) An introduction to stereochemical restraints. *Acta Crystallogr D Biol Crystallogr*, **63**, 58–61.
- Evans, P. R.** (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D Biol Crystallogr*, **67**, 282–292.
- Evans, P. R. y Murshudov, G. N.** (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr*, **69**, 1204–1214.
- Ewald, P. P.** (1969) Introduction to the dynamical theory of X-ray diffraction. *Acta Cryst.*, **A25**, 103–108.
- Falagas, M. E., Makris, G. C., Dimopoulos, G. y Matthaiou, D. K.** (2008) Heteroresistance: a concern of increasing clinical significance? *Clin Microbiol Infect*, **14**, 101–104.
- Fernández-Tornero, C., López, R., García, E., Giménez-Gallego, G. y Romero, A.** (2001) A novel solenoid fold in the cell wall anchoring domain of the pneumococcal virulence factor LytA. *Nat Struct Biol*, **8**, 1020–1024.
- Filipe, S. R., Pinho, M. G. y Tomasz, A.** (2000) Characterization of the *murMN* operon involved in the synthesis of branched peptidoglycan peptides in *Streptococcus pneumoniae*. *J Biol Chem*, **275**, 27768–27774.
- Filipe, S. R. y Tomasz, A.** (2000) Inhibition of the expression of penicillin resistance in *Streptococcus pneumoniae* by inactivation of cell wall mucopeptide branching genes. *Proc Natl Acad Sci U S A*, **97**, 4891–4896.
- Fischer, H. y Tomasz, A.** (1985) Peptidoglycan cross-linking and teichoic acid attachment in *Streptococcus pneumoniae*. *J Bacteriol*, **163**, 46–54.
- Fischer, W.** (1988) Physiology of lipoteichoic acids in bacteria. *Adv Microb Physiol*, **29**, 233–302.
-



- 
- Fischer, W.** (1997) Pneumococcal lipoteichoic and teichoic acid. *Microb Drug Resist*, **3**, 309–325.
- Fischer, W.** (2000) Phosphocholine of pneumococcal teichoic acids: role in bacterial physiology and pneumococcal infection. *Res Microbiol*, **151**, 421–427.
- Folkvardsen, D. B., Thomsen, V. Ø., Rigouts, L., Rasmussen, E. M., Bang, D., Bernaerts, G., Werngren, J., Toro, J. C., Hoffner, S., Hillemann, D. y Svensson, E.** (2013) Rifampin heteroresistance in *Mycobacterium tuberculosis* cultures as detected by phenotypic and genotypic drug susceptibility test methods. *J Clin Microbiol*, **51**, 4220–4222.
- Fraenkel, A.** (1886) Weitere Beitrage zur Lehre von den Mikroccoccen der genuinen fibrinosen Pneumonie. *Zeitschrift für Klinische Medizin*, **11**, 437–458.
- Friedel, G.** (1913) Sur les symétries cristallines que peut révéler la diffraction des rayons Röntgen. *Compt Rend Acad Sci*, **157**, 1533–1536.
- Frolet, C., Beniazza, M., Roux, L., Gallet, B., Noirclerc-Savoye, M., Vernet, T. y Di Guilmi, A. M.** (2010) New adhesin functions of surface-exposed pneumococcal proteins. *BMC Microbiol*, **10**, 190.
- Fulde, M., Bernardo-García, N., Rohde, M., Nachtigall, N., Frank, R., Preissner, K. T., Klett, J., Morreale, A., Chhatwal, G. S., Hermoso, J. A. y Bergmann, S.** (2014) Pneumococcal phosphoglycerate kinase interacts with plasminogen and its tissue activator. *Thromb Haemost*, **111**, 401–416.
- Garbe, J. y Collin, M.** (2012) Bacterial hydrolysis of host glycoproteins - powerful protein modification and efficient nutrient acquisition. *J Innate Immun*, **4**, 121–131.
- García, J. L., Sánchez-Beato, A. R., Medrano, F. J. y López, R.** (1998) Versatility of choline-binding domain. *Microb Drug Resist*, **4**, 25–36.
- García, P., García, J. L., García, E., Sánchez-Puelles, J. M. y López, R.** (1990) Modular organization of the lytic enzymes of *Streptococcus pneumoniae* and its bacteriophages. *Gene*, **86**, 81–88.
- Garcia-Bustos, J. y Tomasz, A.** (1990) A biological price of antibiotic resistance: major changes in the peptidoglycan structure of penicillin-resistant pneumococci. *Proc Natl Acad Sci U S A*, **87**, 5415–5419.
- Garcia-Bustos, J. F., Chait, B. T. y Tomasz, A.** (1987) Structure of the peptide network of pneumococcal peptidoglycan. *J Biol Chem*, **262**, 15400–15405.
- Garman, E. y Murray, J. W.** (2003) Heavy-atom derivatization. *Acta Crystallogr D Biol Crystallogr*, **59**, 1903–1913.
- Garman, E. F.** (2010) Radiation damage in macromolecular crystallography: what is it and why should we care? *Acta Crystallogr D Biol Crystallogr*, **66**, 339–351.
- Garman, E. F. y Mitchell, E. P.** (1996) Glycerol concentrations required for cryoprotection of 50 typical protein crystallization solutions. *J. Appl. Cryst.*, **29**, 584–587.
- Garman, E. F. y Owen, R. L.** (2006) Cryocooling and radiation damage in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*, **62**, 32–47.
-

- 
- Giacovazzo, C.** (2013) *Phasing in crystallography: a modern perspective*. Oxford Science Publications.
- Giacovazzo, C., Monaco, H. L., Artioli, G., Viterbo, D., Milanesio, M., Ferraris, G., Gilli, G., Gilli, P., Zanotti, G. y Catti, M.** (2011) *Fundamentals of crystallography*. Oxford Science Publications.
- Gisch, N., Kohler, T., Ulmer, A. J., Müthing, J., Pribyl, T., Fischer, K., Lindner, B., Hammerschmidt, S. y Zähringer, U.** (2013) Structural reevaluation of *Streptococcus pneumoniae* Lipoteichoic acid and new insights into its immunostimulatory potency. *J Biol Chem*, **288**, 15654–15667.
- Glaser, L. y Lindsay, B.** (1974) The synthesis of lipoteichoic acid carrier. *Biochem Biophys Res Commun*, **59**, 1131–1136.
- Goebel, W. F. y Adams, M. H.** (1943) The immunological properties of the heterophile antigen and somatic polysaccharide of pneumococcus. *J Exp Med*, **77**, 435–449.
- Goehring, N. W. y Beckwith, J.** (2005) Diverse paths to midcell: assembly of the bacterial cell division machinery. *Curr Biol*, **15**, R514–R526.
- Goffin, C. y Ghuysen, J. M.** (1998) Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs. *Microbiol Mol Biol Rev*, **62**, 1079–1093.
- Gosink, K. K., Mann, E. R., Guglielmo, C., Tuomanen, E. I. y Masure, H. R.** (2000) Role of novel choline binding proteins in virulence of *Streptococcus pneumoniae*. *Infect Immun*, **68**, 5690–5695.
- Granville, V., Krivanek, M. y Rasson, J. P.** (1994) Simulated annealing: a proof of convergence. *IEEE TPAMI*, **16**, 652–656.
- Green, D. W., Ingram, V. M. y Perutz, M. F.** (1954) The structure of haemoglobin. IV. Sign determination by the isomorphous replacement method. *Proc R Soc London*, **A**, 287–307.
- Griffith, F.** (1928) The Significance of Pneumococcal Types. *J Hyg (Lond)*, **27**, 113–159.
- Grosse-Kunstleve, R. W. y Adams, P. D.** (2003) Substructure search procedures for macromolecular structures. *Acta Crystallogr D Biol Crystallogr*, **59**, 1966–1973.
- Gutiérrez-Fernández, J., Vaquero, M. E., Prieto, A., Barriuso, J., Martínez, M. J. y Hermoso, J. A.** (2014) Crystal structures of *Ophiostoma piceae* sterol esterase: Structural insights into activation mechanism and product release. *J Struct Biol*, **187**, 215–222.
- Hakenbeck, R., Grebe, T., Zähler, D. y Stock, J. B.** (1999)  $\beta$ -lactam resistance in *Streptococcus pneumoniae*: penicillin-binding proteins and non-penicillin-binding proteins. *Mol Microbiol*, **33**, 673–678.
- Hammerschmidt, S.** (2006) Adherence molecules of pathogenic pneumococci. *Curr Opin Microbiol*, **9**, 12–20.
- Hammerschmidt, S., Talay, S. R., Brandtzaeg, P. y Chhatwal, G. S.** (1997) SpsA, a novel pneumococcal surface protein with specific binding to secretory immunoglobulin A and secretory component. *Mol Microbiol*, **25**, 1113–1124.
-

- 
- Han, H., Yang, Y., Olesen, S. H., Becker, A., Betzi, S. y Schönbrunn, E. (2010) The fungal product terreic acid is a covalent inhibitor of the bacterial cell wall biosynthetic enzyme UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA) . *Biochemistry*, **49**, 4276–4282.
- Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., Carter, 3rd, H. L., Deng, S.-J. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R. y Shewchuk, L. M. (2007) Crystallization of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr*, **63**, 72–79.
- Hendlin, D., Stapley, E. O., Jackson, M., Wallick, H., Miller, A. K., Wolf, F. J., Miller, T. W., Chalet, L., Kahan, F. M., Foltz, E. L., Woodruff, H. B., Mata, J. M., Hernandez, S. y Mochales, S. (1969) Phosphonomycin, a new antibiotic produced by strains of streptomyces. *Science*, **166**, 122–123.
- Hendrickson, W. A. (1991) Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science*, **254**, 51–58.
- Hermoso, J. A., García, J. L. y García, P. (2007) Taking aim on bacterial pathogens: from phage therapy to enzybiotics. *Curr Opin Microbiol*, **10**, 461–472.
- Hermoso, J. A., Lagartera, L., González, A., Stelter, M., García, P., Martínez-Ripoll, M., García, J. L. y Menéndez, M. (2005) Insights into pneumococcal pathogenesis from the crystal structure of the modular teichoic acid phosphorylcholine esterase Pce. *Nat Struct Mol Biol*, **12**, 533–538.
- Hillig, R. C. y Renault, L. (2006) Detecting and overcoming hemihedral twinning during the MIR structure determination of Rna1p. *Acta Crystallogr D Biol Crystallogr*, **62**, 750–765.
- Hirschfield, G. M. y Pepys, M. B. (2003) C-reactive protein and cardiovascular disease: new insights from an old molecule. *QJM*, **96**, 793–807.
- Hofmann, K. y Stoffel, W. (1993) TMbase - A database of membrane spanning proteins segments. *Biol Chem Hoppe-Seyler*, **374**, 166.
- Holm, L. y Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res*, **38**, W545–W549.
- Holmes, A. R., McNab, R., Millsap, K. W., Rohde, M., Hammerschmidt, S., Mawdsley, J. L. y Jenkinson, H. F. (2001) The *pavA* gene of *Streptococcus pneumoniae* encodes a fibronectin-binding protein that is essential for virulence. *Mol Microbiol*, **41**, 1395–1408.
- Hoskins, J., Alborn, Jr, W., Arnold, J., Blaszczyk, L. C., Burgett, S., DeHoff, B. S., Estrem, S. T., Fritz, L., Fu, D. J., Fuller, W., Geringer, C., Gilmour, R., Glass, J. S., Khoja, H., Kraft, A. R., Lagace, R. E., LeBlanc, D. J., Lee, L. N., Lefkowitz, E. J., Lu, J., Matsushima, P., McAhren, S. M., McHenney, M., McLeaster, K., Mundy, C. W., Nicas, T. I., Norris, F. H., O’Gara, M., Peery, R. B., Robertson, G. T., Rockey, P., Sun, P. M., Winkler, M. E., Yang, Y., Young-Bellido, M., Zhao, G., Zook, C. A., Baltz, R. H., Jaskunas, S. R., Rostek,
-

- 
- Jr, P., Skatrud, P. L. y Glass, J. I. (2001) Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J Bacteriol*, **183**, 5709–5717.
- Hoyland, C. N., Aldridge, C., Cleverley, R. M., Duchêne, M.-C., Minasov, G., Onopriyenko, O., Sidiq, K., Stogios, P. J., Anderson, W. F., Daniel, R. A., Savchenko, A., Vollmer, W. y Lewis, R. J. (2014) Structure of the LdcB LD-Carboxypeptidase Reveals the Molecular Basis of Peptidoglycan Recognition. *Structure*, **22**, 949–960.
- Isupov, M. N. y Lebedev, A. A. (2008) NCS-constrained exhaustive search using oligomeric models. *Acta Crystallogr D Biol Crystallogr*, **64**, 90–98.
- Jackson, S. G., Zhang, F., Chindemi, P., Junop, M. S. y Berti, P. J. (2009) Evidence of kinetic control of ligand binding and staged product release in MurA (enolpyruvyl UDP-GlcNAc synthase)-catalyzed reactions. *Biochemistry*, **48**, 11715–11723.
- Joosten, R. P., Womack, T., Vriend, G. y Bricogne, G. (2009) Re-refinement from deposited X-ray data can deliver improved models for most PDB entries. *Acta Crystallogr D Biol Crystallogr*, **65**, 176–185.
- Kabsch, W. (2010a) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr*, **66**, 133–144.
- Kabsch, W. (2010b) XDS. *Acta Crystallogr D Biol Crystallogr*, **66**, 125–132.
- Karle, I. L. y Karle, J. (1981) Anomalous dispersion of sulfur in quinidine sulfate, (C(20)H(25)N(2)O(2))(2)SO(4).2H(2)O: Implications for structure analysis. *Proc Natl Acad Sci U S A*, **78**, 5938–5941.
- Karplus, P. A. y Diederichs, K. (2012) Linking crystallographic model and data quality. *Science*, **336**, 1030–1033.
- Keegan, R. M. y Winn, M. D. (2008) MrBUMP: an automated pipeline for molecular replacement. *Acta Crystallogr D Biol Crystallogr*, **64**, 119–124.
- Keller, P. A., Henrick, K., McNeil, P., Moodie, S. y Barton, G. J. (1998) Deposition of macromolecular structures. *Acta Crystallogr D Biol Crystallogr*, **54**, 1105–1108.
- Kim, D. H., Lees, W. J., Kempell, K. E., Lane, W. S., Duncan, K. y Walsh, C. T. (1996) Characterization of a Cys115 to Asp substitution in the *Escherichia coli* cell wall biosynthetic enzyme UDP-GlcNAc enolpyruvyl transferase (MurA) that confers resistance to inactivation by the antibiotic fosfomycin. *Biochemistry*, **35**, 4923–4928.
- Kirkpatrick, S., Gelatt, Jr, C. y Vecchi, M. P. (1983) Optimization by simulated annealing. *Science*, **220**, 671–680.
- Klein, R. A., Hartmann, R., Egge, H., Behr, T. y Fischer, W. (1994) The aqueous solution structure of the tetrasaccharide-ribitol repeat-unit from the lipoteichoic acid of *Streptococcus pneumoniae* strain R6 determined using a combination of NMR spectroscopy and computer calculations. *Carbohydr Res*, **256**, 189–222.
- Kleywegt, G. J. (2000) Validation of protein crystal structures. *Acta Crystallogr D Biol Crystallogr*, **56**, 249–265.
-

- 
- Korndörfer, I. P., Kanitz, A., Danzer, J., Zimmer, M., Loessner, M. J. y Skerra, A. (2008) Structural analysis of the L-alanoyl-D-glutamate endopeptidase domain of *Listeria* bacteriophage endolysin Ply500 reveals a new member of the LAS peptidase family. *Acta Crystallogr D Biol Crystallogr*, **64**, 644–650.
- Kovacs-Simon, A., Titball, R. W. y Michell, S. L. (2011) Lipoproteins of bacterial pathogens. *Infect Immun*, **79**, 548–561.
- Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
- Laitinen, H. y Tomasz, A. (1990) Changes in composition of peptidoglycan during maturation of the cell wall in pneumococci. *J Bacteriol*, **172**, 5961–5967.
- Langer, G., Cohen, S. X., Lamzin, V. S. y Perrakis, A. (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc*, **3**, 1171–1179.
- Lara, B., Rico, A. I., Petruzzelli, S., Santona, A., Dumas, J., Biton, J., Vicente, M., Mingorance, J. y Massidda, O. (2005) Cell division in cocci: localization and properties of the *Streptococcus pneumoniae* FtsA protein. *Mol Microbiol*, **55**, 699–711.
- Laskowski, R. A. y Swindells, M. B. (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model*, **51**, 2778–2786.
- Leng, J. y Salmon, J.-B. (2009) Microfluidic crystallization. *Lab Chip*, **9**, 24–34.
- Leslie, A. G. W. (2006) The integration of macromolecular diffraction data. *Acta Crystallogr D Biol Crystallogr*, **62**, 48–57.
- Linding, R., Jensen, L. J., Diella, F., Bork, P., Gibson, T. J. y Russell, R. B. (2003) Protein disorder prediction: implications for structural proteomics. *Structure*, **11**, 1453–1459.
- Löffling, J., Vimberg, V., Battig, P. y Henriques-Normark, B. (2011) Cellular interactions by LPxTG-anchored pneumococcal adhesins and their streptococcal homologues. *Cell Microbiol*, **13**, 186–197.
- Long, F., Vagin, A. A., Young, P. y Murshudov, G. N. (2008) BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr*, **64**, 125–132.
- López, R. y García, E. (2004) Recent trends on the molecular biology of pneumococcal capsules, lytic enzymes, and bacteriophage. *FEMS Microbiol Rev*, **28**, 553–580.
- Lopez, R., Garcia, E., Garcia, P., Ronda, C. y Tomasz, A. (1982) Choline-containing bacteriophage receptors in *Streptococcus pneumoniae*. *J Bacteriol*, **151**, 1581–1590.
- Lovering, A. L., Safadi, S. S. y Strynadka, N. C. J. (2012) Structural perspective of peptidoglycan biosynthesis and assembly. *Annu Rev Biochem*, **81**, 451–478.
- Luft, J. R., Wolfley, J. R. y Snell, E. H. (2011) What's in a drop? Correlating observations and outcomes to guide macromolecular crystallization experiments. *Cryst Growth Des*, **11**, 651–663.
- Ma, J. C. y Dougherty, D. A. (1997) The cation- $\pi$  interaction. *Chem Rev*, **97**, 1303–1324.
-

- 
- MacRae, I. J. y Doudna, J. A.** (2007) An unusual case of pseudo-merohedral twinning in orthorhombic crystals of Dicer. *Acta Crystallogr D Biol Crystallogr*, **63**, 993–999.
- Madigan, M. T., Martinko, J. M. y Parker, J.** (2004) *Brock. Biología de los microorganismos*. Pearson Prentice Hall.
- Mann, B., Orihuela, C., Antikainen, J., Gao, G., Sublett, J., Korhonen, T. K. y Tuomanen, E.** (2006) Multifunctional role of choline binding protein G in pneumococcal pathogenesis. *Infect Immun*, **74**, 821–829.
- Manzano, C., Contreras-Martel, C., El Mortaji, L., Izoré, T., Fenel, D., Vernet, T., Schoehn, G., Di Guilmi, A. M. y Dessen, A.** (2008) Sortase-mediated pilus fiber biogenesis in *Streptococcus pneumoniae*. *Structure*, **16**, 1838–1848.
- Massidda, O., Anderluzzi, D., Friedli, L. y Feger, G.** (1998) Unconventional organization of the division and cell wall gene cluster of *Streptococcus pneumoniae*. *Microbiology*, **144** ( Pt 11), 3069–3078.
- Massidda, O., Nováková, L. y Vollmer, W.** (2013) From models to pathogens: how much have we learned about *Streptococcus pneumoniae* cell division? *Environ Microbiol*, **15**, 3133–3157.
- Matsushashi, M., Wachi, M. y Ishino, F.** (1990) Machinery for cell growth and division: penicillin-binding proteins and other proteins. *Res Microbiol*, **141**, 89–103.
- Matthews, B. W.** (1968) Solvent content of protein crystals. *J Mol Biol*, **33**, 491–497.
- Mazmanian, S. K., Liu, G., Jensen, E. R., Lenoy, E. y Schneewind, O.** (2000) *Staphylococcus aureus* sortase mutants defective in the display of surface proteins and in the pathogenesis of animal infections. *Proc Natl Acad Sci U S A*, **97**, 5510–5515.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. y Read, R. J.** (2007) Phaser crystallographic software. *J Appl Crystallogr*, **40**, 658–674.
- McCullers, J. A. y Tuomanen, E. I.** (2001) Molecular pathogenesis of pneumococcal pneumonia. *Front Biosci*, **6**, D877–D889.
- McPherson, A.** (2009) *Introduction to macromolecular crystallography*. John Wiley & Sons.
- Mellroth, P., Sandalova, T., Kikhney, A., Vilaplana, F., Heseck, D., Lee, M., Mobashery, S., Normark, S., Svergun, D., Henriques-Normark, B. y Achour, A.** (2014) Structural and functional insights into peptidoglycan access for the lytic amidase LytA of *Streptococcus pneumoniae*. *MBio*, **5**, e01120–e01113.
- Meroueh, S. O., Bencze, K. Z., Heseck, D., Lee, M., Fisher, J. F., Stemmler, T. L. y Mobashery, S.** (2006) Three-dimensional structure of the bacterial cell wall peptidoglycan. *Proc Natl Acad Sci U S A*, **103**, 4404–4409.
- Merritt, E. A.** (1999) Expanding the model: anisotropic displacement parameters in protein structure refinement. *Acta Crystallogr D Biol Crystallogr*, **55**, 1109–1117.
- Merritt, E. A.** (2012) To B or not to B: a question of resolution? *Acta Crystallogr D Biol Crystallogr*, **68**, 468–477.
-

- 
- Meziane-Cherif, D., Stogios, P. J., Evdokimova, E., Savchenko, A. y Courvalin, P.** (2014) Structural basis for the evolution of vancomycin resistance D,D-peptidases. *Proc Natl Acad Sci U S A*, **111**, 5872–5877.
- Mingorance, J., Tamames, J. y Vicente, M.** (2004) Genomic channeling in bacterial cell division. *J Mol Recognit*, **17**, 481–487.
- Mitchell, T. J.** (2003) The pathogenesis of streptococcal infections: from tooth decay to meningitis. *Nat Rev Microbiol*, **1**, 219–230.
- Mizyed, S., Oddone, A., Byczynski, B., Hughes, D. W. y Berti, P. J.** (2005) UDP-N-acetylmuramic acid (UDP-MurNAc) is a potent inhibitor of MurA (enolpyruvyl-UDP-GlcNAc synthase). *Biochemistry*, **44**, 4011–4017.
- Mizyed, S., Wright, J. E. I., Byczynski, B. y Berti, P. J.** (2003) Identification of the catalytic residues of AroA (Enolpyruvylshikimate 3-phosphate synthase) using partitioning analysis. *Biochemistry*, **42**, 6986–6995.
- Mold, C., Rodgers, C. P., Kaplan, R. L. y Gewurz, H.** (1982) Binding of human C-reactive protein to bacteria. *Infect Immun*, **38**, 392–395.
- Molina, R.** (2009). *Caracterización estructural de la proteína modular CbpF de Streptococcus pneumoniae. Aplicación de complejos de Gd para la obtención de fases en cristalografía macromolecular*. Tesis Doctoral, Universidad Complutense de Madrid.
- Molina, R., González, A., Stelter, M., Pérez-Dorado, I., Kahn, R., Morales, M., Moscoso, M., Campuzano, S., Campillo, N. E., Mobashery, S., García, J. L., García, P. y Hermoso, J. A.** (2009) Crystal structure of CbpF, a bifunctional choline-binding protein and autolysis regulator from *Streptococcus pneumoniae*. *EMBO Rep*, **10**, 246–251.
- Morand, B. y Mühlemann, K.** (2007) Heteroresistance to penicillin in *Streptococcus pneumoniae*. *Proc Natl Acad Sci U S A*, **104**, 14098–14103.
- Morlot, C., Noirclerc-Savoye, M., Zapun, A., Dideberg, O. y Vernet, T.** (2004) The D,D-carboxypeptidase PBP3 organizes the division process of *Streptococcus pneumoniae*. *Mol Microbiol*, **51**, 1641–1648.
- Morlot, C., Pernot, L., Le Gouellec, A., Di Guilmi, A. M., Vernet, T., Dideberg, O. y Dessen, A.** (2005) Crystal structure of a peptidoglycan synthesis regulatory factor (PBP3) from *Streptococcus pneumoniae*. *J Biol Chem*, **280**, 15984–15991.
- Morlot, C., Zapun, A., Dideberg, O. y Vernet, T.** (2003) Growth and division of *Streptococcus pneumoniae*: localization of the high molecular weight penicillin-binding proteins during the cell cycle. *Mol Microbiol*, **50**, 845–855.
- Morris, A. L., MacArthur, M. W., Hutchinson, E. G. y Thornton, J. M.** (1992) Stereochemical quality of protein structure coordinates. *Proteins*, **12**, 345–364.
- Mosser, J. L. y Tomasz, A.** (1970) Choline-containing teichoic acid as a structural component of pneumococcal cell wall and its role in sensitivity to lysis by an autolytic enzyme. *J Biol Chem*, **245**, 287–298.
-

- 
- Murakami, K. y Tomasz, A.** (1989) Involvement of multiple genetic determinants in high-level methicillin resistance in *Staphylococcus aureus*. *J Bacteriol*, **171**, 874–879.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. y Vagin, A. A.** (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr*, **67**, 355–367.
- Musher, D. M.** (1992) Infections caused by *Streptococcus pneumoniae*: clinical spectrum, pathogenesis, immunity, and treatment. *Clin Infect Dis*, **14**, 801–807.
- Nagaoka, S., Murata, S., Kimura, K., Mori, T. y Hojo, K.** (2010) Antimicrobial activity of sodium citrate against *Streptococcus pneumoniae* and several oral bacteria. *Lett Appl Microbiol*, **51**, 546–551.
- Neeleman, C., Geelen, S. P., Aerts, P. C., Daha, M. R., Mollnes, T. E., Roord, J. J., Posthuma, G., van Dijk, H. y Fleer, A.** (1999) Resistance to both complement activation and phagocytosis in type 3 pneumococci is mediated by the binding of complement regulatory protein factor H. *Infect Immun*, **67**, 4517–4524.
- Ness, S. R., de Graaff, R. A. G., Abrahams, J. P. y Pannu, N. S.** (2004) CRANK: new methods for automated macromolecular crystal structure solution. *Structure*, **12**, 1753–1761.
- Neve, H., Zenz, K. I., Desiere, F., Koch, A., Heller, K. J. y Brüßow, H.** (1998) Comparison of the lysogeny modules from the temperate *Streptococcus thermophilus* bacteriophages TP-J34 and Sf21: implications for the modular theory of phage evolution. *Virology*, **241**, 61–72.
- Ng, J. D., Gavira, J. A. y García-Ruíz, J. M.** (2003) Protein crystallization by capillary counterdiffusion for applied crystallographic structure determination. *J Struct Biol*, **142**, 218–231.
- Nodari, C. S., Ribeiro, V. B. y Barth, A. L.** (2015) Imipenem heteroresistance: high prevalence among Enterobacteriaceae *Klebsiella pneumoniae* carbapenemase producers. *J Med Microbiol*, **64**, 124–126.
- Noirclerc-Savoie, M., Le Gouëllec, A., Morlot, C., Dideberg, O., Vernet, T. y Zapun, A.** (2005) In vitro reconstitution of a trimeric complex of DivIB, DivIC and FtsL, and their transient co-localization at the division site in *Streptococcus pneumoniae*. *Mol Microbiol*, **55**, 413–424.
- Novak, R. y Tuomanen, E.** (1999) Pathogenesis of pneumococcal pneumonia. *Semin Respir Infect*, **14**, 209–217.
- Oudega, B., Clark, D., Stegehuis, F., Majoor, M. J. y Luirink, J.** (1993) A lipoprotein signal peptide plus a cysteine residue at the amino-terminal end of the periplasmic protein beta-lactamase is sufficient for its lipid modification, processing and membrane localization in *Escherichia coli*. *FEMS Microbiol Lett*, **108**, 353–359.
- Painter, J. y Merritt, E. A.** (2006) Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr*, **62**, 439–450.
-



- 
- Pannu, N. S., McCoy, A. J. y Read, R. J.** (2003) Application of the complex multivariate normal distribution to crystallographic methods with insights into multiple isomorphous replacement phasing. *Acta Crystallogr D Biol Crystallogr*, **59**, 1801–1808.
- Pannu, N. S. y Read, R. J.** (2004) The application of multivariate statistical techniques improves single-wavelength anomalous diffraction phasing. *Acta Crystallogr D Biol Crystallogr*, **60**, 22–27.
- Pannu, N. S., Waterreus, W. J., Skubák, P., Sikharulidze, I., Abrahams, J. P. y de Graaff, R. A. G.** (2011) Recent advances in the CRANK software suite for experimental phasing. *Acta Crystallogr D Biol Crystallogr*, **67**, 331–337.
- Pape, T. y Schneider, T. R.** (2004) HKL2MAP: a graphical user interface for phasing with SHELX programs. *J Appl Cryst*, **37**, 843–844.
- Parsons, S.** (2003) Introduction to twinning. *Acta Crystallogr D Biol Crystallogr*, **59**, 1995–2003.
- Pasteur, L.** (1881) Note sur la maladie nouvelle provoquée par la salive d'un enfant mort de la rage. *Bulletin de l'Académie de Médecine (Paris)*, **1881**, 94–103.
- Pastoret, S., Fraipont, C., den Blaauwen, T., Wolf, B., Aarsman, M. E. G., Piette, A., Thomas, A., Brasseur, R. y Nguyen-Distèche, M.** (2004) Functional analysis of the cell division protein FtsW of *Escherichia coli*. *J Bacteriol*, **186**, 8370–8379.
- Paterson, N. G., Riboldi-Tunicliffe, A., Mitchell, T. J. y Isaacs, N. W.** (2006) Overexpression, purification and crystallization of a choline-binding protein CbpI from *Streptococcus pneumoniae*. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, **62**, 672–675.
- Pérez-Dorado, I., Galan-Bartual, S. y Hermoso, J. A.** (2012) Pneumococcal surface proteins: when the whole is greater than the sum of its parts. *Mol Oral Microbiol*, **27**, 221–245.
- Pérez-Dorado, I., González, A., Morales, M., Sanles, R., Striker, W., Vollmer, W., Mobashery, S., García, J. L., Martínez-Ripoll, M., García, P. y Hermoso, J. A.** (2010) Insights into pneumococcal fratricide from the crystal structures of the modular killing factor LytC. *Nat Struct Mol Biol*, **17**, 576–581.
- Pflugrath, J. W.** (1999) The finer things in X-ray diffraction data collection. *Acta Crystallogr D Biol Crystallogr*, **55**, 1718–1725.
- Porth, C. M.** (2007) *Fisiopatología. Salud-enfermedad: un enfoque conceptual*. Editorial Médica Panamericana.
- Prescott, L. M., Harley, J. y Klein, D.** (2002) *Microbiology*. The McGraw Hill.
- Pribyl, T., Moche, M., Dreisbach, A., Bijlsma, J. J. E., Saleh, M., Abdullah, M. R., Hecker, M., van Dijk, J. M., Becher, D. y Hammerschmidt, S.** (2014) Influence of impaired lipoprotein biogenesis on surface and exoproteome of *Streptococcus pneumoniae*. *J Proteome Res*, **13**, 650–667.
- Rane, L. y Subbarow, Y.** (1940) Nutritional Requirements of the Pneumococcus: I. Growth Factors for Types I, II, V, VII, VIII. *J Bacteriol*, **40**, 695–704.
-

- 
- Ravelli, R. B. G. y Garman, E. F. (2006) Radiation damage in macromolecular cryocrystallography. *Curr Opin Struct Biol*, **16**, 624–629.
- Raz, R. (2012) Fosfomicin: an old–new antibiotic. *Clin Microbiol Infect*, **18**, 4–7.
- Rezwan, M., Grau, T., Tschumi, A. y Sander, P. (2007) Lipoprotein synthesis in mycobacteria. *Microbiology*, **153**, 652–658.
- Riesenfeld-Orn, I., Wolpe, S., Garcia-Bustos, J. F., Hoffmann, M. K. y Tuomanen, E. (1989) Production of interleukin-1 but not tumor necrosis factor by human monocytes stimulated with pneumococcal cell surface components. *Infect Immun*, **57**, 1890–1893.
- Rigden, D. J., Jedrzejewski, M. J. y Galperin, M. Y. (2003) An extracellular calcium-binding domain in bacteria with a distant relationship to EF-hands. *FEMS Microbiol Lett*, **221**, 103–110.
- Robertson, J. M. (1935) An X-ray study of the structure of the phthalocyanines. Part I. The metal-free, nickel, copper, and platinum compounds. *J Chem Soc*, **1**, 615–621.
- Rodríguez, D. D., Grosse, C., Himmel, S., González, C., de Ilarduya, I. M., Becker, S., Sheldrick, G. M. y Usón, I. (2009) Crystallographic *ab initio* protein structure solution below atomic resolution. *Nat Methods*, **6**, 651–653.
- Rosenow, C., Ryan, P., Weiser, J. N., Johnson, S., Fontan, P., Ortqvist, A. y Masure, H. R. (1997) Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. *Mol Microbiol*, **25**, 819–829.
- Rossmann, M. G. (2001) Molecular replacement–historical background. *Acta Crystallogr D Biol Crystallogr*, **57**, 1360–1366.
- Rossmann, M. G. y Blow, D. M. (1962) The detection of sub-units within the crystallographic asymmetric unit. *Acta Cryst*, **15**, 24–31.
- Rossmann, M. G. y van Beek, C. G. (1999) Data processing. *Acta Crystallographica Section D*, **D55**, 1631 – 1640.
- Roversi, P., Blanc, E., Johnson, S. y Lea, S. M. (2012) Tetartohedral twinning could happen to you too. *Acta Crystallogr D Biol Crystallogr*, **68**, 418–424.
- Rubinson, K. A., Ladner, J. E., Tordova, M. y Gilliland, G. L. (2000) Cryosalts: suppression of ice formation in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*, **56**, 996–1001.
- Rudolph, M. G., Kelker, M. S., Schneider, T. R., Yeates, T. O., Oseroff, V., Heidary, D. K., Jennings, P. A. y Wilson, I. A. (2003) Use of multiple anomalous dispersion to phase highly merohedrally twinned crystals of interleukin-1 $\beta$ . *Acta Crystallogr D Biol Crystallogr*, **59**, 290–298.
- Rupp, B. (2009) *Biomolecular crystallography: principles, practice and application to structural biology*. Garland Science.
- Saleh, M., Bartual, S. G., Abdullah, M. R., Jensch, I., Asmat, T. M., Petruschka, L., Pribyl, T., Gellert, M., Lillig, C. H., Antelmann, H., Hermoso, J. A.
-

- 
- y Hammerschmidt, S.** (2013) Molecular architecture of *Streptococcus pneumoniae* surface thioredoxin-fold lipoproteins crucial for extracellular oxidative stress resistance and maintenance of virulence. *EMBO Mol Med*, **5**, 1852–1870.
- Sánchez-Puelles, J. M., Sanz, J. M., García, J. L. y García, E.** (1990) Cloning and expression of gene fragments encoding the choline-binding domain of pneumococcal murein hydrolases. *Gene*, **89**, 69–75.
- Scapin, G.** (2013) Molecular replacement then and now. *Acta Crystallogr D Biol Crystallogr*, **69**, 2266–2275.
- Schleifer, K. H. y Kandler, O.** (1972) Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev*, **36**, 407–477.
- Schönbrunn, E., Sack, S., Eschenburg, S., Perrakis, A., Krekel, F., Amrhein, N. y Mandelkow, E.** (1996) Crystal structure of UDP-N-acetylglucosamine enolpyruvyltransferase, the target of the antibiotic fosfomycin. *Structure*, **4**, 1065–1075.
- Seo, H. S., Cartee, R. T., Pritchard, D. G. y Nahm, M. H.** (2008) A new model of pneumococcal lipoteichoic acid structure resolves biochemical, biosynthetic, and serologic inconsistencies of the current model. *J Bacteriol*, **190**, 2379–2387.
- Severin, A., Figueiredo, A. M. y Tomasz, A.** (1996) Separation of abnormal cell wall composition from penicillin resistance through genetic transformation of *Streptococcus pneumoniae*. *J Bacteriol*, **178**, 1788–1792.
- Severin, A., Horne, D. y Tomasz, A.** (1997) Autolysis and cell wall degradation in a choline-independent strain of *Streptococcus pneumoniae*. *Microb Drug Resist*, **3**, 391–400.
- Severin, A. y Tomasz, A.** (1996) Naturally occurring peptidoglycan variants of *Streptococcus pneumoniae*. *J Bacteriol*, **178**, 168–174.
- Severin, A., Vaz Pato, M. V., Sá Figueiredo, A. M. y Tomasz, A.** (1995) Drastic changes in the peptidoglycan composition of penicillin resistant laboratory mutants of *Streptococcus pneumoniae*. *FEMS Microbiol Lett*, **130**, 31–35.
- Shankar, K. y Bard, R. C.** (1952) The effect of metallic ions on the growth and morphology of *Clostridium perfringens*. *J Bacteriol*, **63**, 279–290.
- Sheldrick, G. M.** (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr*, **66**, 479–485.
- Shockman, G. D., Daneo-Moore, L., Kariyama, R. y Massidda, O.** (1996) Bacterial walls, peptidoglycan hydrolases, autolysins, and autolysis. *Microb Drug Resist*, **2**, 95–98.
- Skarzynski, T., Kim, D. H., Lees, W. J., Walsh, C. T. y Duncan, K.** (1998) Stereochemical course of enzymatic enolpyruvyl transfer and catalytic conformation of the active site revealed by the crystal structure of the fluorinated analogue of the reaction tetrahedral intermediate bound to the active site of the C115A mutant of MurA. *Biochemistry*, **37**, 2572–2577.
- Skarzynski, T., Mistry, A., Wonacott, A., Hutchinson, S. E., Kelly, V. A. y Duncan, K.** (1996) Structure of UDP-N-acetylglucosamine enolpyruvyl transferase, an
-

---

enzyme essential for the synthesis of bacterial peptidoglycan, complexed with substrate UDP-N-acetylglucosamine and the drug fosfomycin. *Structure*, **4**, 1465–1474.

**Skov-Sørensen, U. B., Blom, J., Birch-Andersen, A. y Henrichsen, J.** (1988) Ultrastructural localization of capsules, cell wall polysaccharide, cell wall proteins, and F antigen in pneumococci. *Infect Immun*, **56**, 1890–1896.

**Sliwiak, J., Jaskolski, M., Dauter, Z., McCoy, A. J. y Read, R. J.** (2014) Likelihood-based molecular-replacement solution for a highly pathological crystal with tetartohedral twinning and sevenfold translational noncrystallographic symmetry. *Acta Crystallogr D Biol Crystallogr*, **70**, 471–480.

**Standish, A. y Morona, R.** (2013) Tyrosine phosphorylation of *Streptococcus pneumoniae* autolysin LytA. In *XI European Meeting on the Molecular Biology of the Pneumococcus*.

**Stein, N.** (2008) CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. *J Appl Cryst*, **41**, 641–643.

**Sternberg, G. M.** (1881) A fatal form of septicaemia in the rabbit, produced by the subcutaneous injection of human saliva. *Annual Reports of the National Board of Health*, **3**, 87–108.

**Strub, M. P., Hoh, F., Sanchez, J. F., Strub, J. M., Böck, A., Aumelas, A. y Dumas, C.** (2003) Selenomethionine and selenocysteine double labeling strategy for crystallographic phasing. *Structure*, **11**, 1359–1367.

**Sugahara, M., Asada, Y., Ayama, H., Ukawa, H., Taka, H. y Kunishima, N.** (2005) Heavy-atom Database System: a tool for the preparation of heavy-atom derivatives of protein crystals based on amino-acid sequence and crystallization conditions. *Acta Crystallogr D Biol Crystallogr*, **61**, 1302–1305.

**Swiatlo, E., Champlin, F. R., Holman, S. C., Wilson, W. W. y Watt, J. M.** (2002) Contribution of choline-binding proteins to cell surface properties of *Streptococcus pneumoniae*. *Infect Immun*, **70**, 412–415.

**Sylvester, D. R., Alvarez, E., Patel, A., Ratnam, K., Kallender, H. y Wallis, N. G.** (2001) Identification and characterization of UDP-N-acetylenolpyruvylglucosamine reductase (MurB) from the Gram-positive pathogen *Streptococcus pneumoniae*. *Biochem J*, **355**, 431–435.

**Taylor, G. L.** (2010) Introduction to phasing. *Acta Crystallogr D Biol Crystallogr*, **66**, 325–338.

**Terwisscha van Scheltinga, A. C., Valegård, K., Hajdu, J. y Andersson, I.** (2003) MIR phasing using merohedrally twinned crystals. *Acta Crystallogr D Biol Crystallogr*, **59**, 2017–2022.

**Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., Heidelberg, J., DeBoy, R. T., Haft, D. H., Dodson, R. J., Durkin, A. S., Gwinn, M., Kolonay, J. F., Nelson, W. C., Peterson, J. D., Umayam, L. A., White, O., Salzberg, S. L., Lewis, M. R., Radune, D., Holtzapple, E., Khouri, H., Wolf, A. M., Utterback, T. R., Hansen, C. L., McDonald, L. A., Feldblyum,**

- 
- T. V., Angiuoli, S., Dickinson, T., Hickey, E. K., Holt, I. E., Loftus, B. J., Yang, F., Smith, H. O., Venter, J. C., Dougherty, B. A., Morrison, D. A., Hollingshead, S. K. y Fraser, C. M.** (2001) Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science*, **293**, 498–506.
- Thouvenin, E. y Hewat, E.** (2000) When two into one won't go: fitting in the presence of steric hindrance and partial occupancy. *Acta Crystallogr D Biol Crystallogr*, **56**, 1350–1357.
- Tickle, I. J., Laskowski, R. A. y Moss, D. S.** (1998) Rfree and the rfree ratio. I. Derivation of expected values of cross-validation residuals used in macromolecular least-squares refinement. *Acta Crystallogr D Biol Crystallogr*, **54**, 547–557.
- Tickle, I. J., Laskowski, R. A. y Moss, D. S.** (2000) Rfree and the rfree ratio. II. Calculation of the expected values and variances of cross-validation statistics in macromolecular least-squares refinement. *Acta Crystallogr D Biol Crystallogr*, **56**, 442–450.
- Tillet, W. S., Goebel, W. F. y Avery, O. T.** (1930) Chemical and immunological properties of a species-specific carbohydrate of pneumococci. *J Exp Med*, **52**, 895–900.
- Tomasz, A.** (1967) Choline in the cell wall of a bacterium: novel type of polymer-linked choline in *Pneumococcus*. *Science*, **157**, 694–697.
- Tomasz, A.** (1968) Biological consequences of the replacement of choline by ethanolamine in the cell wall of *Pneumococcus*: choline formation, loss of transformability, and loss of autolysis. *Proc Natl Acad Sci U S A*, **59**, 86–93.
- Tomasz, A.** (1970) Cellular metabolism in genetic transformation of pneumococci: requirement for protein synthesis during induction of competence. *J Bacteriol*, **101**, 860–871.
- Tong, L.** (2001) How to take advantage of non-crystallographic symmetry in molecular replacement: 'locked' rotation and translation functions. *Acta Crystallogr D Biol Crystallogr*, **57**, 1383–1389.
- Tu, A. H., Fulgham, R. L., McCrory, M. A., Briles, D. E. y Szalai, A. J.** (1999) Pneumococcal surface protein A inhibits complement activation by *Streptococcus pneumoniae*. *Infect Immun*, **67**, 4720–4724.
- Tuomanen, E., Liu, H., Hengstler, B., Zak, O. y Tomasz, A.** (1985) The induction of meningeal inflammation by components of the pneumococcal cell wall. *J Infect Dis*, **151**, 859–868.
- Tuomanen, E., Rich, R. y Zak, O.** (1987) Induction of pulmonary inflammation by components of the pneumococcal cell surface. *Am Rev Respir Dis*, **135**, 869–874.
- Tuomanen, E., Tomasz, A., Hengstler, B. y Zak, O.** (1985) The relative role of bacterial cell wall and capsule in the induction of inflammation in pneumococcal meningitis. *J Infect Dis*, **151**, 535–540.
- Usón, I. y Sheldrick, G. M.** (1999) Advances in direct methods for protein crystallography. *Curr Opin Struct Biol*, **9**, 643–648.
- Vagin, A. y Teplyakov, A.** (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr*, **66**, 22–25.
-

- 
- van Heijenoort, J.** (1998) Assembly of the monomer unit of bacterial peptidoglycan. *Cell Mol Life Sci*, **54**, 300–304.
- Vellieux, F. M. y Read, R. J.** (1997) Noncrystallographic symmetry averaging in phase refinement and extension. *Methods Enzymol*, **277**, 18–53.
- Vollmer, W. y Tomasz, A.** (2000) The *pgdA* gene encodes for a peptidoglycan N-acetylglucosamine deacetylase in *Streptococcus pneumoniae*. *J Biol Chem*, **275**, 20496–20501.
- Ward, J. J., McGuffin, L. J., Bryson, K., Buxton, B. F. y Jones, D. T.** (2004) The DISOPRED server for the prediction of protein disorder. *Bioinformatics*, **20**, 2138–2139.
- Webb, M.** (1951) The influence of magnesium on cell division. The effect of magnesium on the growth of bacteria in chemically-defined media of varying complexity. *J Gen Microbiol*, **5**, 485–495.
- Weber, B., Ehlert, K., Diehl, A., Reichmann, P., Labischinski, H. y Hakenbeck, R.** (2000) The *fib* locus in *Streptococcus pneumoniae* is required for peptidoglycan crosslinking and PBP-mediated beta-lactam resistance. *FEMS Microbiol Lett*, **188**, 81–85.
- Weiss, M. S.** (2001) Global indicators of X-ray data quality. *J. Appl. Cryst.*, **34**, 130–135.
- Weiss, M. S., Sicker, T. y Hilgenfeld, R.** (2001) Soft X-rays, high redundancy, and proper scaling: a new procedure for automated protein structure determination via SAS. *Structure*, **9**, 771–777.
- Welsh, A. H., Lindenmayer, D. B. y Donnelly, C. F.** (2013) Fitting and interpreting occupancy models. *PLoS One*, **8**, e52015.
- Wiegels, T. y Lamzin, V. S.** (2012) Use of noncrystallographic symmetry for automated model building at medium to low resolution. *Acta Crystallogr D Biol Crystallogr*, **68**, 446–453.
- Wilson, J. J., Matsushita, O., Okabe, A. y Sakon, J.** (2003) A bacterial collagen-binding domain with novel calcium-binding motif controls domain orientation. *EMBO J*, **22**, 1743–1752.
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. y Wilson, K. S.** (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr*, **67**, 235–242.
- Winn, M. D., Isupov, M. N. y Murshudov, G. N.** (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr*, **57**, 122–133.
- Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K. y Uversky, V. N.** (2010) PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim Biophys Acta*, **1804**, 996–1010.
-

- 
- Yang, C., Pflugrath, J. W., Courville, D. A., Stence, C. N. y Ferrara, J. D.** (2003) Away from the edge: SAD phasing from the sulfur anomalous signal measured in-house with chromium radiation. *Acta Crystallogr D Biol Crystallogr*, **59**, 1943–1957.
- Yeates, T. O.** (1988) Simple statistics for intensity data from twinned specimens. *Acta Crystallogr A*, **44** ( Pt 2), 142–144.
- Yeates, T. O.** (1997) Detecting and overcoming crystal twinning. *Methods Enzymol*, **276**, 344–358.
- Yeates, T. O. y Yu, F.** (2008) Equations for determining tetartohedral twin fractions. *Acta Crystallogr D Biol Crystallogr*, **64**, 1158–1164.
- Yoshikawa, H., Hirano, A., Arakawa, T. y Shiraki, K.** (2012) Mechanistic insights into protein precipitation by alcohol. *Int J Biol Macromol*, **50**, 865–871.
- Yother, J., Leopold, K., White, J. y Fischer, W.** (1998) Generation and properties of a *Streptococcus pneumoniae* mutant which does not require choline or analogs for growth. *J Bacteriol*, **180**, 2093–2101.
- Zavascki, A. P., Falci, D. R., da Silva, R. C. F., Dalarosa, M. G., Ribeiro, V. B., Rozales, F. P., Luz, D. I., Magagnin, C. M., Vieira, F. J., Sampaio, J. M. y Barth, A. L.** (2014) Heteroresistance to carbapenems in New Delhi metallo- $\beta$ -lactamase-1-producing isolates: a challenge for detection? *Infect Control Hosp Epidemiol*, **35**, 751–752.
- Zhang, Z., Wang, Y., Pang, Y. y Liu, C.** (2014) Comparison of different drug susceptibility test methods to detect rifampin heteroresistance in *Mycobacterium tuberculosis*. *Antimicrob Agents Chemother*, **58**, 5632–5635.
- Zhu, J.-Y., Yang, Y., Han, H., Betzi, S., Olesen, S. H., Marsilio, F. y Schönbrunn, E.** (2012) Functional consequence of covalent reaction of phosphoenolpyruvate with UDP-N-acetylglucosamine 1-carboxyvinyltransferase (MurA). *J Biol Chem*, **287**, 12657–12667.
-





---

---

## ARTÍCULOS DE INVESTIGACIÓN

---

---



**Hansjürg, E., Gutiérrez-Fernández, J., Flückiger, C., Martínez-Ripoll, M., Mühlemann, K., Hermoso, J. A., Hilty, M. y Hathaway, L. J.** (2013) Heteroresistance to fosfomycin is predominant in *Streptococcus pneumoniae* and depends on the *murA1* gene. *Antimicrob Agents Chemother*, **57**(6), 2801–2808.

# Heteroresistance to Fosfomycin Is Predominant in *Streptococcus pneumoniae* and Depends on the *murA1* Gene

Hansjürg Engel,<sup>a,b</sup> Javier Gutiérrez-Fernández,<sup>c</sup> Christine Flückiger,<sup>a</sup> Martín Martínez-Ripoll,<sup>c</sup> Kathrin Mühlemann,<sup>a,d,†</sup> Juan A. Hermoso,<sup>c</sup> Markus Hilty,<sup>a,d</sup> Lucy J. Hathaway<sup>a</sup>

Institute for Infectious Diseases, Faculty of Medicine, University of Bern, Bern, Switzerland<sup>a</sup>; Graduate School for Cellular and Biomedical Sciences, University of Bern, Bern, Switzerland<sup>b</sup>; Department of Crystallography and Structural Biology, Instituto de Química-Física Rocasolano, CSIC, Madrid, Spain<sup>c</sup>; Department of Infectious Diseases, University Hospital, Bern, Switzerland<sup>d</sup>

Fosfomycin targets the first step of peptidoglycan biosynthesis in *Streptococcus pneumoniae* catalyzed by UDP-*N*-acetylglucosamine enolpyruvyltransferase (MurA1). We investigated whether heteroresistance to fosfomycin occurs in *S. pneumoniae*. We found that of 11 strains tested, all but 1 (Hungary<sup>19A</sup>) displayed heteroresistance and that deletion of *murA1* abolished heteroresistance. Hungary<sup>19A</sup> differs from the other strains by a single amino acid substitution in MurA1 (Ala<sub>364</sub>Thr). To test whether this substitution is responsible for the lack of heteroresistance, it was introduced into strain D39. The heteroresistance phenotype of strain D39 was not changed. Furthermore, no relevant structural differences between the MurA1 crystal structures of heteroresistant strain D39 and nonheteroresistant strain Hungary<sup>19A</sup> were found. Our results reveal that heteroresistance to fosfomycin is the predominant phenotype of *S. pneumoniae* and that MurA1 is required for heteroresistance to fosfomycin but is not the only factor involved. The findings provide a caveat for any future use of fosfomycin in the treatment of pneumococcal infections.

Pneumococci have responded to antimicrobial treatment by acquiring resistance to antibiotics, particularly penicillin (1). However, although knowledge of classical resistance mechanisms is increasing (2, 3), understanding of the penicillin heteroresistance phenomenon described in pneumococci remains elusive (4).

Heteroresistance is the ability of a clonal population to grow one or several subpopulations at a frequency of 10<sup>-7</sup> to 10<sup>-3</sup> in the presence of a higher antibiotic concentration than that predicted to be effective by measurement of the MIC (4). This phenomenon may allow some bacteria to grow at higher antibiotic concentrations without the fitness cost required by a genetically fixed resistance mechanism. Furthermore, heteroresistance allows some bacteria to survive antimicrobial treatment while at the same time making conventional resistance testing unreliable (4, 5). Heteroresistance to different antibiotics has been described in pathogens of various species (6–9), but even for the most extensively studied (heteroresistance to methicillin, oxacillin, and vancomycin in staphylococci [10–16]), the mechanism is unknown. Understanding heterogeneity between single cells is complicated by the fact that conventional assays of microbial populations consider an averaged value of thousands or millions of cells in a sample (17).

Fosfomycin (18) is not currently used in the treatment of pneumococcal infections, but in this era of multidrug resistance, studies of this antibiotic may be valuable to inform decisions on future antibiotic use (19). Fosfomycin is a drug of particular interest for its synergy with benzylpenicillin, by decreasing the production of penicillin-binding proteins. This is beneficial in the treatment of infections with  $\beta$ -lactam-resistant pneumococci (20). Combinations of other antibiotics with fosfomycin are used successfully in the treatment of infections with other organisms, such as *Staphylococcus aureus* (21). The target of fosfomycin is MurA (UDP-*N*-acetylglucosamine enolpyruvyltransferase, EC 2.5.1.7), an enzyme catalyzing the

first step of the peptidoglycan biosynthetic pathway of Gram-positive and Gram-negative microorganisms. MurA transfers the enolpyruvyl group of phosphoenolpyruvate (PEP) to UDP-*N*-acetylglucosamine (UDP-GlcNAc), producing enolpyruvyl UDP-*N*-acetylglucosamine (EP-UDP-GlcNAc) (22, 23). Gram-positive bacteria have two genes, *murA1* and *murA2*, both encoding active MurA enzymes with similar catalytic parameters and conserved major structural features that are both inhibited by fosfomycin (24).

Mutagenesis experiments have revealed that fosfomycin covalently binds to Cys115 of MurA, suggesting that this amino acid plays an important role in the catalytic mechanism (25, 26). The Cys115 residue has been suggested to be the general acid catalyst, protonating C-3 of PEP prior to nucleophilic attack by UDP-GlcNAc to form the tetrahedral intermediate (27, 28), although other experiments indicate that a different amino acid (Asp) plays this role (29). Recent studies have shown that MurA forms dormant complexes with UDP-*N*-acetylmuramic acid (UDP-MurNAc), which is the product of MurB, the second enzyme partici-

Received 1 February 2013 Returned for modification 21 February 2013

Accepted 4 April 2013

Published ahead of print 9 April 2013

Address correspondence to Lucy J. Hathaway, lucy.hathaway@ifik.unibe.ch.

† Deceased 1 November 2012.

M.H. and L.J.H. contributed equally to this work.

We dedicate this work to the memory of Kathrin Mühlemann, who designed and guided research for this study, encouraging us to overcome difficulties of any kind. Her expertise and friendship are sorely missed.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AAC.00223-13>.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AAC.00223-13

---

## Heteroresistance to fosfomycin in *S. pneumoniae* - Supplementary Material

### Methods

**Knock-out mutants.** For the construction of *murA1* knock-out mutants, *murA1* upstream and downstream flanking regions were amplified and fused to the bicistronic cassette Janus. Fast Start Taq DNA polymerase (Roche Molecular Biochemicals, Rotkreuz, Switzerland) was used for amplification if not stated otherwise. For primers and PCR conditions see Table S1. The resulting fragment was introduced into the recipient strains 106.44, Taiwan<sup>23F</sup>, Hungary<sup>19A</sup> and D39 as described before (34) using Columbia sheep blood agar (CSBA) plates containing 500µg/ml kanamycin to select transformants (if not stated otherwise all chemicals from Sigma, USA). Insertion was confirmed by PCR.

**Hungary::*murA1*<sub>D39</sub> and D39::*murA1*<sub>D39</sub> back-transformant.** For the construction of mutants D39::*murA1*<sub>D39</sub> back-transformant and Hungary::*murA1*<sub>D39</sub> the *murA1* of D39 and its up- and downstream flanking regions were amplified and fused to a spectinomycin cassette obtained from pR412 plasmid. The obtained construct was used to transform D39*murA1*::Janus strains in order to obtain the D39::*murA1*<sub>D39</sub> back-transformant and Hungary*murA1*::Janus strain to produce strain Hungary::*murA1*<sub>D39</sub>. Transformants were selected on CSBA plates containing 200µg/ml spectinomycin under anaerobic conditions.

**Cloning *murA1*<sub>D39</sub>-spectinomycin construct for mutagenesis.** For the insertion of *murA1*-spectinomycin construct into pGEX-6p-1 plasmid the *murA1*-spectinomycin construct was amplified from D39::*murA1*<sub>D39</sub> back-transformant and ligated into pGEM-T Easy vector (pGEM-T Easy Vector System, Promega, Switzerland) according to the manufacturer's protocol. Inserts of plasmids obtained from minipreps (Wizard Plus SV Minipreps DNA purification System, Promega, USA) were

---

sequenced as described earlier (36). Deletion of guanine at position 1249 of the insert was detected.

The plasmid was digested with *ScaI* restriction enzyme (all restriction enzymes from New England Biolabs, USA), then *NotI*-HF. The *murA1*-spectinomycin-construct was isolated from a 1% agarose gel. In parallel pGEX-6p-1 plasmid (GE Healthcare, Sweden) was linearized with *NotI*-HF, then dephosphorylated with Shrimp Alkaline Phosphatase (SAP). The *murA1*-spectinomycin construct was inserted using T4 DNA ligase (New England Biolabs, USA) overnight at 16°C. Transformation of competent *E. coli* SURE cells with 10 µl of the ligation reaction, was performed at 42°C for 2 min. Transformants were selected on Luria-Bertani-Agar (LBA) plates containing 100 µg/ml ampicillin and 200 µg/ml spectinomycin. Plasmid pGEX::*murA1*<sub>D39</sub>-spect-1 was isolated and insert checked by sequencing as described before (36).

**Insert G<sub>1090</sub>A into *murA1*<sub>D39</sub>-spectinomycin construct for D39::*murA1*<sub>Hungary19A</sub> and Hungary::*murA1*<sub>Hungary19A</sub> back-transformant.** For the construction of D39::*murA1*<sub>Hungary19A</sub> and Hungary::*murA1*<sub>Hungary19A</sub> back-transformant mutation G<sub>1090</sub>A was introduced into pGEX::*murA1*<sub>D39</sub>-spect-1 by QuikChange II Site-Directed Mutagenesis Kit (Stratagene, USA), using *E. coli* SURE cells for cloning thus obtaining pGEX::*murA1*<sub>Hungary19A</sub>-spect. Deletion G<sub>1249</sub> was cured by a second mutagenesis reaction using Phusion High Fidelity Polymerase (Fisher Scientific, Switzerland). The product obtained was purified and template plasmid was digested with *DpnI* restriction enzyme overnight at 37°C, then ligated at 4°C for 3 days with T4 DNA ligase. The construct was amplified with Phusion High-Fidelity polymerase. The amplified construct containing mutation G<sub>1090</sub>A, characteristic of *murA1*<sub>Hungary19A</sub> and re-inserted G<sub>1249</sub> was isolated from 1% agarose gel and used to transform D39*murA1*::Janus to get the *murA1* switch-mutant and Hungary*murA1*::Janus to

---

obtain the Hungary::*murA*<sub>1Hungary19A</sub> back-transformants as described before. The presence of inserted constructs was confirmed by sequencing.

**Hungary::*murA*<sub>1</sub>+*murB*<sub>D39</sub> and D39::*murA*<sub>1</sub>+*murB*<sub>Hungary19A</sub> mutants.** For the construction of Hungary::*murA*<sub>1</sub>+*murB*<sub>D39</sub> and D39::*murA*<sub>1</sub>+*murB*<sub>Hungary19A</sub> mutants in order to switch *murB* sequences between strains D39 and Hungary<sup>19A</sup> the *murB*<sub>D39</sub> gene was amplified. It was then fused to the kanamycin resistance gene amplified from the Janus cassette, as was the downstream flanking region of *murB*<sub>Hungary19A</sub>. Amplification was performed using iProof polymerase (Bio-Rad, Switzerland), fragments were fused as described elsewhere (35) with Phusion polymerase. The *murB*<sub>D39</sub>-kanamycin construct was then transformed into Hungary::*murA*<sub>D39</sub> as described above and transformants selected on CSBA plates containing 200 µg/ml kanamycin. A second construct was put together with *murB*<sub>Hungary19A</sub>. *MurB*<sub>Hungary19A</sub> lacks the last 15 amino acids (Ser<sub>302</sub> – Arg<sub>316</sub>) compared to *MurB*<sub>D39</sub>. The *murB*<sub>Hungary19A</sub>-kanamycin construct was transformed into D39::*murA*<sub>1Hungary19A</sub>. Replacement of genes was confirmed by sequencing.

**Recombinant protein expression.** The *murA*<sub>1</sub> sequences of strains D39 and Hungary<sup>19A</sup> were inserted into pGEX-6p-1 plasmid at its XhoI and BamHI sites. SAP was used for dephosphorylation. Ligation was performed with T4 DNA ligase at 4°C for 3 days. *E. coli* SURE cells were transformed and transformants selected on LBA with 100 µg/ml ampicillin. Plasmids were isolated with Promega Wizard and used to transform competent *E. coli* BL21 cells for expression. The *murA*<sub>1</sub> sequence was confirmed in isolated plasmids by sequencing.

Single clones of *E. coli* BL21::pGEX-*murA*<sub>1D39</sub> and *E. coli* BL21::pGEX-*murA*<sub>1Hungary19A</sub> were grown in LB broth with 100 µg/ml ampicillin to OD<sub>600</sub> 0.5 at 37°C at 150 rpm then protein expression was induced with 0.1 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG) overnight at 20°C and 150 rpm. Cells were pelleted (10

---

min, 4000 g) and lysed on ice under native conditions with 25 ml of lysis buffer (50mM NaH<sub>2</sub>PO<sub>4</sub>, 300mM NaCl, 10mM Imidazol, 1 mg/ml lysozyme, pH 8.0) (Roche, Switzerland) for 30 min. Proteases were inhibited by addition of 1 µg/ml aprotinine, 1 µg/ml leupeptine, 1 µg/ml pepstatine and 1 mM phenylmethanesulfonyl fluoride (PMSF). Protein purification was performed on Glutathione Sepharose 4 Fast Flow (GE Healthcare, Sweden) as suggested by the manufacturer using PreScission Protease (GE Healthcare, Sweden) to cleave recombinant proteins from the GST-Tag. Recombinant MurA1 was separated from the beads with Qiagen Polypropylene columns. The buffer was exchanged to protein stocking buffer (10mM Tris-HCl, pH 7.5) on Amicon Ultra-4 filter device (Millipore, Ireland). Protein samples were analyzed by SDS-PAGE and were >95 % of the expected protein. Identity of recombinant proteins was confirmed by MS peptide fingerprinting.



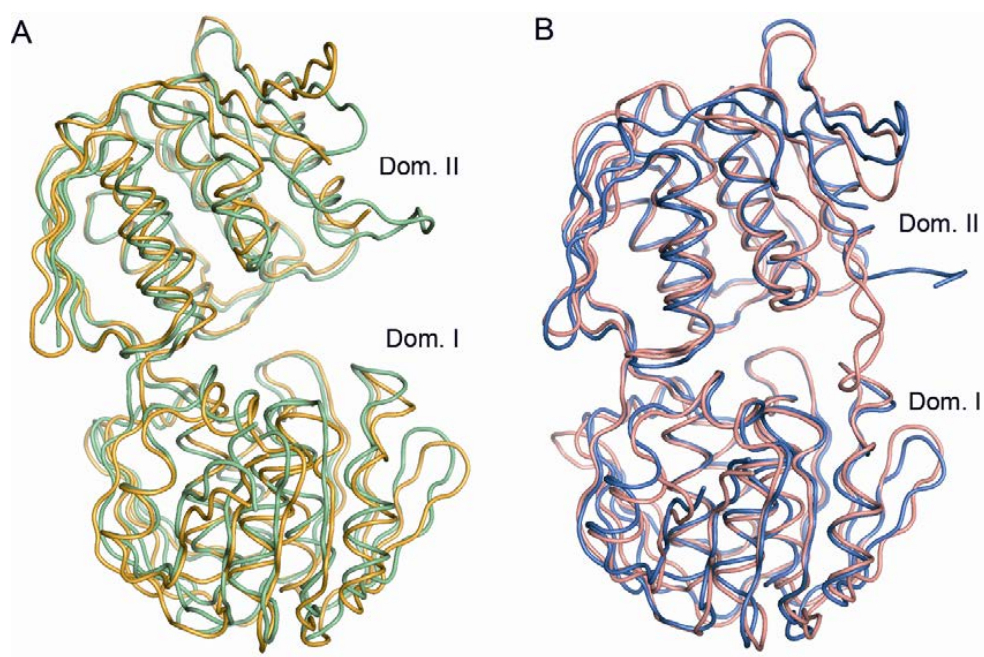
Table S1 — Primers used in construction of mutants & probes

Primer	Sequence 5'→3'	Purpose	PCR conditions:Annealing temperature / Extension time
<b><i>murA1</i>-Janus construct</b>			
KoMurA1_f2	5'-AAACTCTAAAGTCCAAAAGGTAGTCGTAA-3'	Amplification upstream <i>murA1</i>	57 °C / 1 min 30 s
DAM406_rev+b3	5'-TCTGGAATAGGCATAGATCTAAGTTTCAATACTCTTT-3'	Amplification upstream <i>murA1</i> , overlap Janus cassette	57 °C / 1 min 30 s
DAM351_rev+f4	5'-ATAAGGAAAGGGCCCTAGACCGTAGAGGTGTTTATGA-3'	Amplification downstream <i>murA1</i> , overlap Janus cassette	60 °C / 1 min 30 s
KoMurA1_b5	5'-CCAGCCAGAACCTCAGAAGAAGTC-3'	Amplification downstream <i>murA1</i>	60 °C / 1 min 30 s
b3_rev+DAM406	5'-AAAGAGTATTGAAAACCTAGATCTATGCCTATTCCAGA-3'	Amplification Janus cassette, overlap upstream <i>murA1</i>	61 °C / 1 min 30 s
f4_rev+DAM351	TCATAAACACCTCTACGGTCTAGGGCCCCCTTTCCTTAT	Amplification of Janus cassette, overlap downstream <i>murA1</i>	61 °C / 1 min 30 s
DAM351	5'-CTAGGGCCCCCTTTCTTATGCTTTTGGAC-3'	Fusion <i>murA1</i> upstream - Janus cassette	55.2 °C / 3 min
KoMurA1_f2A	5'-GATGTAGCACTTATCCCGTAGAGT-3'	Fusion <i>murA1</i> downstream - <i>murA1</i> upstream + Janus cassette	55.7 °C / 3 min
KoMurA1_b5A	5'-CGAAATGCTACGAACACTGAC-3'	Fusion <i>murA1</i> downstream - <i>murA1</i> upstream + Janus cassette	55.7 °C / 3 min
rpsL_mitte_b1	5'-GAATTACGAAGGGCTGAGTTAGGTTT-3'	Control insertion Janus cassette, with KoMurA1_f2A	55 °C / 2 min
Kan_start_f1	5'-AGGAAATAATAAATGGCTAAAATGAGAAT-3'	Control insertion Janus cassette, with KoMurA1_b5B	50 °C / 2 min
<b><i>murA1</i>-spectinomycin construct</b>			
MurA1_B8rev+spect_f1	5'-AGACTTGTTGAGGATTAACCTAGTGGATCCCCGTT-3'	Amplification SPT cassette, overlap upstream <i>murA1</i>	58 °C / 1 min 30 s
spect_b1+KoMurA1_f4rev	5'-CATAAACACCTCTACGGTATAGTTCCTTCAAGAGC-3'	Amplification SPT cassette, overlap <i>murA1</i> downstream flanking region	58 °C / 1 min 30 s
spect_b1rev+KoMurA1+KoMurA1_f4	5'-GCTCTTGAAGGAACTATACCGTAGAGGTGTTTATG-3'	Amplification <i>murA1</i> downstream, overlap SPT cassette	58 °C / 1 min 30 s
KoMurA1_b6	5'-ATCATGAACCTGAAATCCCCCTGTA-3'	Amplification <i>murA1</i> downstream, overlap SPT cassette	58 °C / 1 min 30 s
KoMurA1_f1	5'-AGTATATCTGATGTCAAAGTCTCGGCG-3'	Amplification <i>murA1</i> and upstream, overlap SPT cassette	59 °C / 2 min 10 s
MurA1_B8+spect_f1rev	5'-AACGGGGGATCCACTAGTTTAATCCTCAACAAGTCT-3'	Amplification <i>murA1</i> and upstream, overlap SPT cassette	59 °C / 2 min 10 s
KoMurA1_b5B	5'-CCGACTTTTCTTGTTGTAACGA-3'	Fusion <i>murA1</i> SPT construct, with KoMurA1_f2A	52.3 °C / 3 min
KoMurA1_f2B	5'-GGTATGTATAAACGCTTCTGTGA-3'	Aplification whole <i>murA1</i> SPT construct, with KoMurA1_b5B	52 °C / 3 min
pGEX_sequencing_F	5'-GGGCTGCGAACGCCAGTTCGGTG-3'	Amplification <i>murA1</i> hungary19A+G1249, seq. pGEX:: <i>murA1</i> D39-spect-1	68 °C / 1 min
pGEX_sequencing_R	5'-CCGGGAGCTGCATGTGTGACAGG-3'	Amplification <i>murA1</i> hungary19A+G1249, seq. pGEX:: <i>murA1</i> D39-spect-1	68 °C / 1 min
<b>Mutagenesis primers</b>			
SDMspC_G1090A_f	5'-GGTGGAAGTATACGTGGGACCAAGTGTAAAGCGACC-3'	Introduce G1090A in pGEX:: <i>murA1</i> D39-spect-1	68 °C / 8 min 30 s
SDMspC_G1090A_r	5'-GGTCCGCTTTAAACACTGGTCCCACGTAATCAGTCCACC-3'	Introduce G1090A in pGEX:: <i>murA1</i> D39-spect-1	68 °C / 8 min 30 s
SDMexp_G1249_f	5'-PGAGGCGGATATAGACTGTTGAGGATTAACCTAGTGGATC-3'	Cure deletion G1249; Primer 5' phosphorylated	65° C / 4 min
SDMexp_G1249_r	5'-P-GATCCACTAGTTTAATCCTCAACAAGTCTAATATCCGCTCC-3'	Cure deletion G1249; Primer 5' phosphorylated	65° C / 4 min
<b><i>murB</i>-kanamycin construct</b>			
KoMurB_f1	5'-CGTCAATACCTGTTTGCTAAGTTCAATGTTG-3'	Amplify <i>murB</i> D39 or <i>murB</i> Hungary19A+ upstream, sequencing	55° C / 40 s
murB_D39_PCR1_r	5'-GCCATTTATTATTTCTTCTCTCTACTCTTGCAAGGAGATAAACCC-3'	Amplify <i>murB</i> D39 + upstream, overlap KAN cassette	55° C / 40 s
murB_D39_PCR2_f	5'-GGTTTACTCCCTGCAAGAGGTAGAGAGGAAGGAAATAATAATGGGC-3'	Amplify KAN cassette, overlap <i>murB</i> D39	55 °C / 20 s
murB_H_PCR2_r	5'-GCAGCAGCTTTTCAAATACAGCTTCTCTAAACAATTCATCCAG-3'	Amplify KAN cassette, overlap <i>murB</i> D39	55 °C / 20 s
murB_H_PCR3_f	5'-CTGGATGAATTGTTTAGGAAAGCTGATTGTAAGCTGCTGC-3'	Amplify downstream <i>murB</i> Hungary19A, overlap KAN cassette	55 °C / 20 s
Ko_murB_b6	5'-GGTTCATTGGCTTCATCCCACC-3'	Amplify downstream <i>murB</i> Hungary19A, overlap KAN cassette	55 °C / 20 s
murB_H_PCR1_r	5'-GCCATTTATTATTTCTTCTCTCTACTTCTGCTTTCAACCAAGATCCGGAC-3'	Amplify <i>murB</i> Hungary19+ upstream, overlap KAN cassette	55° C / 40 s
murB_H_PCR2_f	5'-GTCCGGATCTGGGTGAAAGCAAGTAGAGAGGAAGGAAATAATAATGGCC-3'	Amplify KAN cassette, overlap <i>murB</i> Hungary19A	55 °C / 20 s
murB_D_PCR2_r	5'-GGGCTCTGTGAGGTCCCACATAAACAAATTCATCCAGTA-3'	Amplify KAN cassette, overlap <i>murB</i> Hungary19A	55 °C / 20 s
murB_D_PCR3_f	5'-CTGATGAATGTTTATAGTGGGACCTGACAGAGCCC-3'	Amplify downstream <i>murB</i> D39, overlap KAN cassette	55 °C / 20 s
Ko_murB_us_F	5'-GGACATGTCATGGATTTTGTCAACAAG-3'	Amplify whole <i>murB</i> KAN construct for sequencing	55 °C / 1 min
Ko_murB_ds_R	5'-CCATTCAATTTCCAAGTTCGTATAGG-3'	Amplify whole <i>murB</i> KAN construct for sequencing	55 °C / 1 min
<b>Recombinant proteins</b>			
pGEX_murA1_For3	5'-GAGGATCCATGAGAAAAATTGTTATCAATGGTGGA-3'	Amplification <i>murA1</i> incl. restriction site <u>Bam</u> HI	51 °C / 2 min
pGEX_murA1_Rev2	5'-AAACTCGAGTTATTAATCCTCAACAAGTCTAATATC-3	Amplification <i>murA1</i> incl. restriction site XhoI	51 °C / 2 min

Primer	Sequence 5'→3'	Purpose	PCR conditions: Annealing temperature / Extension time
<b>Sequencing primers</b>			
T7_FOR	5'-TAA TAC GAC TCA CTA TAG GG-3'	Seq. pGEM-T::murA1D39-SPT and pGEX::murA1D39-spect-1	-
MurA1_For	5'-ATGAGAAAAATTGTTATCAATGGTGGA-3'	Seq. pGEM-T::murA1D39-SPT and pGEX::murA1D39-SPT-1	-
SeqSpecC_1306_F	5'-AGGCCGTTTTGGTGAAGCGACAGT-3'	Seq. pGEM-T::murA1D39-SPT and pGEX::murA1D39-SPT-1	-
SpecC_1921_F	5'-GCAACAACCGCTTACCCCTCTTTT-3'	Seq. pGEM-T::murA1D39-SPT and pGEX::murA1D39-SPT-1	-
SeqSpecC_2938_F	5'-GCGGGAATGCAGTGGCTGAATCT-3'	Seq. pGEM-T::murA1D39-SPT and pGEX::murA1D39-SPT-1	-
SeqSpecC_3292_F	5'-TTATTTCCCAACGCAGGCAAGTC-3'	Seq. pGEM-T::murA1D39-SPT and pGEX::murA1D39-SPT-1	-
SeqSpecC_3943_R	5'-GCCCAGCCAGAACCTCAGAAGAAGT-3'	Seq. pGEM-T::murA1D39-SPT and pGEX::murA1D39-SPT-1	-
murA1_Rev	5'-TTAATCCTCAACAAGTCTAATATCCGC-3'	Seq. pGEM-T::murA1D39-SPT and pGEX::murA1D39-SPT-1	-
SP6_REV	5'-TATTAGTGAACTATAG-3'	Seq. pGEM-T::murA1D39-SPT and pGEX::murA1D39-SPT-1	-
MurB_for	5'-GTGACATGTCTGTAAGAGAAAAATGCTTG-3'	Seq. murB-KAN construct	-
MurB_for2	5'-CTCACATCTTGCACTCTTGTAAAGTCTTG-3'	Seq. murB-KAN construct	-
murB_Rev	5'-GTAAACCACCTGCATACATCTTCGC-3'	Seq. murB-KAN construct	-
KoMurB_b5	5'-CACGGGGTTGGTTGATGATGACACG-3'	Seq. murB-KAN construct	-
<b>Gene expression studies</b>			
MURA1_F	5'-CTCGTCCGATTGACTTACACCTTAA-3'	Measure gene expression murA1	-
MURA1_R	5'-ACTTCATGTTATCTCCCTCGTAGCT-3'	Measure gene expression murA1	-
MURA1_M	5'-CCATAGCTTCAAACGCC-3'	Measure gene expression murA1 - FAM probe	-
murA2-igt2F	5'-CCAGGTGGTTGTACGATTGGTA-3'	Measure gene expression murA2	-
murA2-igt2R	5'-CCCATAGCTTCCAGACCTTTCAAA-3'	Measure gene expression murA2	-
murA2-igt2M1	5'-CCGTCCTATTGATCTTCA-3'	Measure gene expression murA2 - FAM probe	-
TaqMan 16S_For	5'-GACGATACATAGCCGACCTGAGA-3'	Measure gene expression 16S	-
TaqMan 16S_Rev	5'-GTAGGAGTCTGGGCGGTGTCT-3'	Measure gene expression 16S	-
16S-igt1M2probe	5'-CCAGTGTGGCCGATC-3'	Measure gene expression 16S - FAM probe	-

Abbreviations: SPT (spectinomycin); KAN (kanamycin); incl. (including); seq. (sequencing)

**Figure S1**



**Figure S1 - Closest structural homologues for MurA1<sub>D39</sub> and for MurA1<sub>Hungary19A</sub>** (A) Structural superimposition of MurA1<sub>D39</sub> (orange) and open conformation of the MurA from *E. cloacae* (green, PDB code 1EJD), rmsd = 1.8 Å. (B) Structural superimposition of MurA1Hungary19A (blue) and the closed conformation of MurA from *B. anthracis* (pink, PDB code 3SG1), rmsd = 1.3 Å.

---

**Figure S2**

	113									122
<i>Streptococcus pneumoniae</i>	P	G	G	C	D	L	G	P	R	P
<i>Enterobacter cloacae</i>	P	G	G	C	A	I	G	A	R	P
<i>Bacillus anthracis</i>	P	G	G	C	A	I	G	S	R	P
<i>Escherichia coli</i>	P	G	G	C	T	I	G	A	R	P
<i>Vibrio fischeri</i>	P	G	G	C	A	I	G	A	R	P
<i>Haemophilus influenzae</i>	P	G	G	C	S	I	G	A	R	P
<i>Listeria monocytogenes</i>	P	G	G	C	A	I	G	S	R	P
<i>Aquifex aeolicus</i>	P	G	G	C	S	I	G	A	R	P
<i>Mycobacterium tuberculosis</i>	P	G	G	D	A	I	G	S	R	P

**Figure S2 - Sequence alignment of the MurA1 flexible loop from several species.** The loop containing catalytic Cys116 presents some differences compared to that of close homologues (highlighted in a box). In *S. pneumoniae* the loop differs in the presence of an Asp residue after Cys116 and one more Pro residue at position 120.



Abdullah, M. R., Gutiérrez-Fernández, J., Pribyl, T., Gisch, N., Saleh, M., Rohde, M., Petruschka, L., Burchhardt, G., Schwudke, D., Hermoso, J. A. y Hammerschmidt, S. (2014) Structure of the pneumococcal L,D-carboxypeptidase DacB and pathophysiological effects of disabled cell wall hydrolases DacA and DacB. *Mol Microbiol*, **93**(6), 1183–1206.

# Structure of the pneumococcal L,D-carboxypeptidase DacB and pathophysiological effects of disabled cell wall hydrolases DacA and DacB

Mohammed R. Abdullah,<sup>1†</sup>  
Javier Gutiérrez-Fernández,<sup>2†</sup> Thomas Pribyl,<sup>1</sup>  
Nicolas Gisch,<sup>3</sup> Malek Saleh,<sup>1</sup> Manfred Rohde,<sup>4</sup>  
Lothar Petruschka,<sup>1</sup> Gerhard Burchhardt,<sup>1</sup>  
Dominik Schwudke,<sup>3</sup> Juan A. Hermoso<sup>2\*\*</sup> and  
Sven Hammerschmidt<sup>1\*</sup>

<sup>1</sup>Department Genetics of Microorganisms, Interfaculty Institute for Genetics and Functional Genomics, Ernst Moritz Arndt University of Greifswald, D-17487 Greifswald, Germany.

<sup>2</sup>Department of Crystallography and Structural Biology, Institute of Physical-Chemistry 'Rocasolano', CSIC, Serrano 119, 28006 Madrid, Spain.

<sup>3</sup>Division of Bioanalytical Chemistry, Research Center Borstel, Leibniz-Center for Medicine and Biosciences, D-23845 Borstel, Germany.

<sup>4</sup>Department of Molecular Infection Biology, Central Facility for Microscopy, Helmholtz Centre for Infection Research, D-38124 Braunschweig, Germany.

## Summary

Bacterial cell wall hydrolases are essential for peptidoglycan turnover and crucial to preserve cell shape. The D,D-carboxypeptidase DacA and L,D-carboxypeptidase DacB of *Streptococcus pneumoniae* function in a sequential manner. Here, we determined the structure of the surface-exposed lipoprotein DacB. The crystal structure of DacB, radically different to that of DacA, contains a mononuclear Zn<sup>2+</sup> catalytic centre located in the middle of a large and fully exposed groove. Two different conformations were found presenting a different arrangement of the active site topology. The critical residues for catalysis and substrate specificity were identified. Loss-of-function of DacA and DacB altered the cell shape and this was consistent with a modified peptidoglycan

peptide composition in *dac* mutants. Contrary, an *lgt* mutant lacking lipoprotein diacylglycerol transferase activity required for proper lipoprotein maturation retained L,D-carboxypeptidase activity and showed an intact murein sacculus. In addition we demonstrated pathophysiological effects of disabled DacA or DacB activities. Real-time bioimaging of intranasal infected mice indicated a substantial attenuation of  $\Delta$ *dacB* and  $\Delta$ *dacA* $\Delta$ *dacB* pneumococci, while  $\Delta$ *dacA* had no significant effect. In addition, uptake of these mutants by professional phagocytes was enhanced, while the adherence to lung epithelial cells was decreased. Thus, structural and functional studies suggest DacA and DacB as optimal drug targets.

## Introduction

The Gram-positive bacterium *Streptococcus pneumoniae*, commonly known as the pneumococcus, colonizes often symptomless as harmless commensal the upper respiratory tract of humans. However, these ellipsoid-shaped facultative anaerobic bacteria can also turn into a serious pathogen causing a variety of respiratory and life-threatening invasive diseases that range from mild middle ear infections (otitis media) to severe community-acquired pneumonia, bacteraemia and meningitis (Cartwright, 2002; Siemieniuk *et al.*, 2011; Gamez and Hammerschmidt, 2012). Despite the availability of vaccines and the use of antibiotics, pneumococci are responsible for > 1.5 million death worldwide and kill more people in the United States and Europe than other vaccine-preventable disease (Atkinson *et al.*, 2009; Gamez and Hammerschmidt, 2012).

In the past three decades, pneumococcal resistance to a variety of antimicrobial agents has rapidly evolved from a rare phenomenon to a worldwide health problem. According to the Centers for Disease Control and Prevention, there were about 37 000 cases of invasive pneumococcal disease in the USA in 2011. In 31% of the cases pneumococci were resistant to one or more antibiotics (CDC, 2013). The spread of drug resistant pneumococci has now reached the level of concern and clinical relevance. Espe-

Accepted 22 July, 2014. For correspondence. \*E-mail sven.hammerschmidt@uni-greifswald.de; Tel. (+49) 3834 864161; Fax (+49) 3834 864172. \*\*E-mail xjuan@iqfr.csic.es; Tel. (+34) 917459538; Fax (+34) 915642431. †These authors contributed equally to this work.

© 2014 John Wiley & Sons Ltd

---

## SUPPORTING INFORMATION

### Structure of the pneumococcal L,D-carboxypeptidase DacB and pathophysiological effects of disabled cell wall hydrolases DacA and DacB

Mohammed R. Abdullah<sup>1‡</sup>, Javier Gutiérrez-Fernández<sup>2‡</sup>, Thomas Pribyl<sup>1</sup>, Nicolas Gisch<sup>3</sup>,  
Malek Saleh<sup>1</sup>, Manfred Rohde<sup>4</sup>, Lothar Petruschka<sup>1</sup>, Gerhard Burchhardt<sup>1</sup>, Dominik  
Schwudke<sup>3</sup>, Juan A. Hermoso<sup>2\*</sup> and Sven Hammerschmidt<sup>1\*</sup>

<sup>1</sup>Department Genetics of Microorganisms, Interfaculty Institute for Genetics and Functional Genomics, Ernst Moritz Arndt University of Greifswald, D-17487 Greifswald, Germany

<sup>2</sup>Department of Crystallography and Structural Biology, Institute of Physical-Chemistry "Rocasolano", CSIC, Serrano 119, 28006-Madrid, Spain

<sup>3</sup>Division of Bioanalytical Chemistry, Research Center Borstel, Leibniz-Center for Medicine and Biosciences, D-23845 Borstel, Germany

<sup>4</sup>Department of Molecular Infection Biology, Central Facility for Microscopy, Helmholtz Centre for Infection Research, D-38124 Braunschweig, Germany

#### Experimental procedures

#### References

Table S1

Table S2

Table S3

Figure S1

Figure S2

Figure S3

Figure S4

Figure S5

Figure S6

Figure S7

Figure S8

Figure S9

Figure S10

Figure S11

Figure S12

Figure S13

Figure S14

Figure S15

Figure S16

Figure S17

Figure S18



---

## Experimental procedures

### *Flow cytometry*

*S. pneumoniae* wild-type and isogenic mutants were cultured in 30 mL THY to A<sub>600</sub> of 0.35 to 0.4 and after sedimentation the bacteria were washed with RPMI1640/1% FBS (fetal bovine serum; PAA Laboratories, Cölbe, Germany) and finally resuspended in 1 mL of RPMI1640/1% FBS. To detect DacA and DacB, respectively, on the surface of pneumococci  $1 \times 10^8$  bacteria were incubated with mouse anti-DacB polyclonal antisera (1:500 dilution in PBS) or antisera of mice injected with PBS and incubated for 30 min at 4°C. To investigate the effect of the loss-of-function of carboxypeptidases on the amounts of capsular polysaccharide, wild-type strain D39 and its isogenic *dac*-mutants were incubated with the serotype 2 specific antiserum (Statens Serum Institute, Denmark). Samples were then washed twice with PBS/0.5% fetal calf serum (FCS) and stained with secondary goat anti-mouse or rabbit IgG coupled Alexa-Fluor-488 (Life Technologies). After 30 min incubation at 4°C bacteria were washed twice with PBS/0.5% and then fixed with 2% paraformaldehyde. Flow cytometry was conducted with the FACS Calibur™ (BD Biosciences, Franklin Lakes, U.S.) and the CellQuestPro Software 6.0. (BD Biosciences) was used for data acquisition, while analysis of the data was performed with the software WinMDI 2.9.

### *Electron microscopy*

#### *Transmission electron microscopy*

For examination of the pneumococcal ultrastructure bacteria were fixed with 2.5% glutaraldehyde, 2% formaldehyde containing 0.5% tannic acid in cacodylate buffer (0.1M cacodylate, 0.01 M CaCl<sub>2</sub>, 0.01 M MgCl<sub>2</sub>, 0.09 M sucrose, pH 6.9) for 1 hour on ice, washed with cacodylate buffer, and further fixed with 1% aqueous osmium for 1 hour at room temperature. Dehydration was achieved with a graded series of acetone (10, 30, 50, 70, 90, 100%) on ice. Samples were then embedded in the epoxy resin Spurr according to described procedures (Spurr, 1969). Ultrathin sections were cut with a diamond knife, picked

---

up with butvar-coated grids, counterstained with uranyl acetate and examined in a TEM 910 transmission electron microscope (Carl Zeiss, Oberkochen) at an acceleration voltage of 80 kV. Images were recorded digitally at calibrated magnifications with a Slow-Scan CCD-Camera (ProScan, 1024x1024, Scheuring, Germany) with ITEM-Software (Olympus Soft Imaging Solutions, Münster, Germany). Contrast and brightness were adjusted with Adobe Photoshop CS5.

#### *Field emission scanning electron microscopy (FESEM)*

For morphological studies pneumococci were fixed as described above, kept at 4°C before washing with cacodylate buffer and TE-buffer (20 mM TRIS, 1 mM EDTA, pH 6,9). Samples were placed on poly-L-lysine coated glass cover slips, allowed to settle for 5 min, fixed with 2% glutaraldehyde at room temperature, and subsequently washed with TE-buffer. Samples were dehydrated in a graded series of acetone (10, 30, 50, 70, 90, 100%) on ice for 15 min for each step. Samples in the 100% acetone step were allowed to reach room temperature before another change in 100% acetone. Samples were then subjected to critical-point drying with liquid CO<sub>2</sub> (CPD 300, Leica). Dried samples were covered with a gold-palladium film by sputter coating (SCD 500, Bal-Tec, Liechtenstein) before examination in a field emission scanning electron microscope Zeiss Merlin using the Everhart Thornley SE detector and the inlens detector in a 25:75 ratio at an acceleration voltage of 5 kV at calibrated magnifications. Contrast and brightness were adjusted with Adobe Photoshop CS5.

#### *Procedure for peptidoglycan-peptide isolation*

Strains were grown in THY-medium [pH 7.4] to the late logarithmic phase ( $OD_{600} \approx 1$ ) and then harvested by centrifugation (8,000 × g, 10 min, r.t.). After washing with citric buffer (50 mM, pH 4.7) cells were resuspended in citric buffer containing 4% SDS and incubated for 20 min at 100 °C. Cells were then stored at -80 °C overnight and lyophilized subsequently. Following an earlier described protocol (Gisch *et al.*, 2013) for pnLTA isolation, pneumococcal cells were resuspended in distilled water prior to extraction, thus restoring the

---

concentration of citric buffer (50 mM, pH 4.7) and SDS (4%). Cells were disrupted in a cell homogenizer (Vibrogen Cell Mill VI 6, Edmund Bühler GmbH) with 0.1 mm glass beads (Carl Roth GmbH). Afterwards, glass beads were removed by centrifugation ( $2,000 \times g$ , 10 min, r.t.) and washed two times with citric buffer. To the combined supernatants SDS (20% solution) was added to reobtain a total concentration of 4%. This solution was incubated for 30 min at 100 °C and centrifuged ( $30,000 \times g$ , 15 min, 4 °C) subsequently. The pellet was washed four times with citric buffer. (The pnLTA containing supernatants were combined and lyophilized) After washing the pellet SDS-free, it is treated with DNase, RNase, and trypsin, followed by boiling in 1% SDS, ultra centrifugation, washing with 8 M LiCl, 10 mM EDTA, acetone and water (following the protocol of (Bui *et al.*, 2012). The resulting pellet containing cell walls was used for LytA treatment. Cell walls were stirred at 37 °C in 50 mM Tris-HCl (pH 7.0) [10 mg/mL] with His-tagged LytA amidase (1 mg/mL) added in three aliquots (0.33 mg LytA/100 mg cell walls each aliquot) after 0, 24 and 48 h for a total incubation period of 72 h. Afterwards the sample was centrifuged ( $25,000 \times g$ , 15 min, 20 °C) and the supernatant was lyophilized. The obtained material was subsequently separated by gel permeations chromatography (GPC) using a BioGel P-30 column (120  $\times$  1.5 cm) with 150 mM NH<sub>4</sub>Ac buffer (pH 4.7) as eluting buffer, as an example the chromatogram for peptides of D39 $\Delta$ cps $\Delta$ dacB is shown in Fig. 4B. This profile is characteristic for all performed P-30 separations (specific fraction numbers can slightly vary). Here, in pool 1 predominantly the dimeric peptides are eluting, in pool 2 the monomeric ones. Such a strict separation was not achieved for all strains. Additionally, in the earlier fractions the tri- to oligomeric peptides are eluting, but have not been used in this study.

#### *Analysis of the peptidoglycan-peptide containing GPC pools by mass spectrometry*

Electrospray Ionization Fourier-Transform Ion Cyclotron Resonance Mass Spectrometry (ESI-FT-ICR-MS) was performed on a 7 Tesla APEX Qe instrument (Bruker Daltonics, USA). Overview mass spectra of the GPC pools (Figs. 4C, 4D, and S4-S8) were recorded in the negative-ion mode, for which the samples were dissolved in a water/propan-2-

---

ol/triethylamine/acetic acid mixture (50:50:0.06:0.02 v/v/v/v). Most spectra were acquired with micro-ESI using a flow rate of 2  $\mu$ l/min with a capillary voltage set to -3.8 kV, the pools of D39 $\Delta$ cps $\Delta$ lgt (Fig. 4D and Fig. S8) were analysed with nano-ESI using the Triversa Nanomate (Advion, USA) as ion source with a spray voltage set to -1.1 kV. Mass spectra were recorded in broadband mode and the mass scale was externally calibrated with glycolipids of known structure. The spectra were charge deconvoluted and mass numbers given refer to the monoisotopic mass of the neutral molecules. Only the relevant sections of the measured survey scans are depicted, respectively.

#### *Separation of PGN-peptides by HPLC and their analysis by ESI-MS(/MS)*

PGN-peptides were separated on a reversed-phase column (Uptisphere Strategy C18-2 3  $\mu$ m, 4.6  $\times$  250 mm, 100 Å (Interchim, France); pre-column: ZORBAX ODS 5  $\mu$ m, 4  $\times$  5 mm (Bischoff, Germany)) with a flow rate of 0.5 ml/min. Elution occurred starting with water/0.1% TFA (solvent A) and acetonitrile/0.1% TFA (solvent B) in a 99:1 (v/v)-mixture for 10 min, followed by a linear 80-min gradient from 1% to 20% solvent B. After a subsequent linear 20-min gradient from 20% to 50% solvent B, the solvent mixture of 50:50 (v/v) was held for 10 min. Finally the column was re-equilibrated during 30 min to 1% solvent B and hold there for additional 10 min prior to the next injection. For recording PGN-peptide elution profiles, 20  $\mu$ g of the respective GPC pools were measured as triplicates on a Dionex ICS-3000 system including autosampler and UltiMate 3000 Diode Array Detector with Chromeleon 6.80 as operating software, representative profiles are shown in figures S9 and S10. For the isolation of separated PGN-peptides similar HPLC runs using 100 to 200  $\mu$ g of the PGN-peptide mixtures have been performed on a Gilson HPLC system including an UV/VIS-151 detector and an FC203B fraction collector (Unipoint 3.30 as operation software), for monomeric PGN-peptide containing GPC pools with slightly altered gradient (starting with 0% solvent B instead of 1%). PGN-peptide fractions detected at 210 nm were collected for ESI-MS(/MS) analysis. For this, samples were dried under nitrogen, dissolved in a small amount of water and measured in a water/propan-2-ol/30 mM ammonium acetate/acetic acid mixture

---

(15:15:1:0.04 v/v/v/v) or water/acetonitrile/formic acid (70:30:0.1 v/v/v), respectively. All peptides were annotated according to retention time and MS/MS interpretation. In figure S11 the MS/MS-analysis of **3** and its glycine-containing variant **3\*** is depicted as an example, proving the glycine to be in position 5 of the stem peptide, what is in agreement with the literature (Bui *et al.*, 2012).

#### *Mouse lung infection model and bioluminescent optical imaging*

Eight weeks old female outbred CD1 mice (Charles River, Sulzfeld, Germany) were infected intranasally with bioluminescent pneumococci as described recently. Briefly, pneumococci were cultured to A<sub>600</sub> of 0.35 in THY supplemented with 10% FBS and the infection dose was adjusted to  $1.0 \times 10^7$  CFU in 25  $\mu$ L for the intranasal route (n=10). Before intranasal infection, mice were anesthetized by intraperitoneal injection of ketamine (Ketanest S; Pfizer Pharma, Karlsruhe, Germany) and xylazine (Rompun<sup>®</sup>; Provet AG, Lyssach, Switzerland). Once anesthetized the animals were scuffed, with the nose held upright, and the bacterial suspension of 25  $\mu$ L was administered intranasally by adding a series of small droplets into the nostrils for the mice to involuntarily inhale. The infection dose was confirmed by determination of the CFU after plating serial dilutions of the infection dose on blood agar plates. Bioluminescent optical imaging using the IVIS<sup>®</sup> Spectrum Imaging System (Caliper Life Sciences, Hopkinton, U.S.) allowed monitoring of pneumococcal dissemination after intranasal infection (Hartel *et al.*, 2011, Jensch *et al.*, 2010, Saleh *et al.*, 2014). At pre-chosen time intervals post-infection mice were imaged for 1 min to monitor dissemination of bioluminescent pneumococci into the lungs. A time series of the captured images was generated and the bioluminescent intensity (BLI) was determined by quantification of the total photon emission using the Living Image<sup>®</sup> 4.1 software package (Caliper Life Sciences).

#### *Phagocytosis experiments*

To determine the rate of phagocytosed wild-type and mutant pneumococci and their intracellular survival in macrophages, phagocytosis experiments with J774A.1 murine

---

macrophages (DSMZ, Braunschweig, Germany) were carried out as described previously (Hartel et al., 2011, Jensch et al., 2010). Briefly, confluent monolayers of J774 cells in 96-well cell culture plates and cultured in RPMI1640/10% FBS (PAA Laboratories) were incubated for 30 min with pneumococci. Extracellular bacteria and non-adherent pneumococci were then killed by the antibiotic protection assay. The infected cells were washed three times with the infection medium and then replaced and incubated for 1 h with RPMI1640 medium containing gentamicin (100 µg/mL) and penicillin G (100 units/mL) to kill non-internalized bacteria. Intracellular pneumococci were recovered by a saponin-mediated lysis (1% w/v) of macrophages (Jensch et al., 2010) and the number of recovered and viable pneumococci (survivors) was determined by quantitative plating of the released intracellular pneumococci on sheep blood agar plates (Oxoid). Experiments were conducted at least three times with four replicate wells. Fluorescence microscopy was performed to visualize attached or ingested pneumococci. Macrophages ( $10^5$ ) were seeded on glass coverslips (12 mm Ø) in wells of a 24-well cell culture plate and infected the following day with pneumococci as described above. Post-infection unbound bacteria were removed and the infected host cells were fixed on the glass coverslips with 3.7% paraformaldehyde) and double immunofluorescence microscopy was carried out as described (Jensch et al., 2010). Image acquisition was performed with a fluorescence microscope (Zeiss Axio-Observer.Z1 with VisiGrid, Coolsnap HQ and Visiview Imaging software (Visitron Systems GmbH, Puchheim, Germany). Each bar in the images represents 20 µm. All experiments were performed at least three times with two or more replicate wells tested for each experimental setup.

#### *Cell culture adherence assays and immunofluorescence microscopy*

Adherence assays were conducted with human epithelial cell lines A549 (ATCC CCI-185; type II pneumocytes). This cell line was cultured and infected as described previously (Bergmann et al., 2009). Briefly, epithelial cells were seeded on glass coverslips (12 mm Ø) or directly in the wells of a 24-well plate (Cellstar; Greiner Bio-One, Kremsmünster, Austria)

---

at a density of  $5 \times 10^4$  cells per well and cultivated to confluent cell layers with approximately  $2 \times 10^5$  cells per well. The cells were washed three times with Dulbecco's modified Eagle's medium containing HEPES (DMEM-HEPES, PAA Laboratories) supplemented with 1% FCS and then infected with pneumococci. In a standardized assay we used a multiplicity of infection (MOI) of 25 bacteria per host cell and infections were carried out for 3 h at 37°C and 5% CO<sub>2</sub>. Thereafter, unbound bacteria were removed by rinsing three times with DMEM-HEPES with 1% FCS. To determine the number of host cell associated bacteria the cells with attached and intracellular bacteria were plated on blood agar plates and the colony forming units were determined. For immunofluorescence microscopy the infected host cells were fixed on the glass coverslips with 3% paraformaldehyde and host cell-bound bacteria were stained using a polyclonal anti-pneumococcal antiserum and secondary goat anti-rabbit IgG coupled Alexa-Fluor-488 (green) (Life Technologies). The infected host cells were fixed on the glass coverslips with 3% paraformaldehyde and host cell-bound bacteria were stained using a polyclonal anti-pneumococcal antiserum and secondary goat anti-rabbit IgG coupled Alexa-Fluor-488 (green) (Life Technologies). Image acquisition was performed with a fluorescence microscope (Zeiss Axio-Observer.Z1 with VisiGrid, Coolsnap HQ and Visiview Imaging software (Visitron Systems GmbH, Puchheim, Germany). Each bar in the images represents 20 μm. All experiments were performed at least three times with two or more replicate wells tested for each experimental setup.

#### *Quantification of pneumococcal invasion*

Antibiotic protection assays were performed to quantify the total number of internalized and recovered pneumococci after the infection experiments, which were carried out under different conditions. The infections were performed for 3 hours as described and thereafter the cell layers were washed thoroughly to remove unbound bacteria. To kill the extracellular, adherent pneumococci, the host cells were incubated for 1 hour with DMEM-HEPES containing 100 μg gentamicin and 100 U penicillin G at 37°C and 5% CO<sub>2</sub>. The intracellular pneumococci were released by a saponin mediated host-cell lysis (1% w/v). The total

---

number of invasive and recovered pneumococci was monitored after plating sample aliquots on blood agar plates, followed by enumeration of CFU. Each experiment was repeated at least three times and results were expressed as mean  $\pm$  SD

## REFERENCES

- Bergmann, S., A. Lang, M. Rohde, V. Agarwal, C. Rennemeier, C. Grashoff, K.T. Preissner and S. Hammerschmidt (2009) Integrin-linked kinase is required for vitronectin-mediated internalization of *Streptococcus pneumoniae* by host cells. *J Cell Sci* **122**: 256-267.
- Bui, N.K., A. Eberhardt, D. Vollmer, T. Kern, C. Bougault, A. Tomasz, J.P. Simorre and W. Vollmer (2012) Isolation and analysis of cell wall components from *Streptococcus pneumoniae*. *Anal Biochem* **421**: 657-666.
- Gisch, N., T. Kohler, A.J. Ulmer, J. Muthing, T. Pribyl, K. Fischer, B. Lindner, S. Hammerschmidt and U. Zahringer (2013) Structural reevaluation of *Streptococcus pneumoniae* Lipoteichoic acid and new insights into its immunostimulatory potency. *J Biol Chem* **288**: 15654-15667.
- Hartel, T., M. Klein, U. Koedel, M. Rohde, L. Petruschka and S. Hammerschmidt (2011) Impact of glutamine transporters on pneumococcal fitness under infection-related conditions. *Infect Immun* **79**: 44-58.
- Jensch, I., G. Gamez, M. Rothe, S. Ebert, M. Fulde, D. Somplatzki, S. Bergmann, L. Petruschka, M. Rohde, R. Nau and S. Hammerschmidt (2010) PavB is a surface-exposed adhesin of *Streptococcus pneumoniae* contributing to nasopharyngeal colonization and airways infections. *Mol Microbiol* **77**: 22-43.
- Saleh, M., M.R. Abdullah, C. Schulz, T. Kohler, T. Pribyl, I. Jensch and S. Hammerschmidt (2014) Following in real time the impact of pneumococcal virulence factors in an acute mouse pneumonia model using bioluminescent bacteria. *J Vis Exp*: e51174.
- Saleh, M., S.G. Bartual, M.R. Abdullah, I. Jensch, T.M. Asmat, L. Petruschka, T. Pribyl, M. Gellert, C.H. Lillig, H. Antelmann, J.A. Hermoso and S. Hammerschmidt (2013) Molecular architecture of *Streptococcus pneumoniae* surface thioredoxin-fold lipoproteins crucial for extracellular oxidative stress resistance and maintenance of virulence. *EMBO Mol Med* **5**: 1852-1870.
- Spurr, A.R. (1969) A low-viscosity epoxy resin embedding medium for electron microscopy. *J Ultrastruct Res* **26**: 31-43.



Table S1: Antibiotic susceptibilities of pneumococcal *dac*-mutants

Antibiotics	MIC (µg/ml)			
	D39	D39Δ <i>dacA</i>	D39Δ <i>dacB</i>	D39Δ <i>dacA</i> Δ <i>dacB</i>
Ampicillin	≤ 0.25	≤ 0.25	≤ 0.25	≤ 0.25
Cefotaxime	≤ 0.12	≤ 0.12	≤ 0.12	≤ 0.12
Ceftriaxone	≤ 0.12	≤ 0.12	≤ 0.12	≤ 0.12
Vancomycin	≤ 0.5	≤ 0.25	≤ 0.25	≤ 0.25

The interpretation of antibiotic susceptibility was carried out according to the European Committee on Antimicrobial Susceptibility Testing (ECAST) standard.

Table S2. Strain and plasmid list.

Strain or plasmid	Serotype and relevant Genotype <sup>a</sup>	Resistance	Source or Reference
<b><i>Streptococcus pneumoniae</i></b>			
D39	2	None	NCTC7466
TIGR4	4	None	Tettelin <i>et al.</i> , 2001
PN111	D39Δ <i>cps</i>	Km <sup>r</sup>	Rennemeier <i>et al.</i> , 2007
PN149	D39/ <i>lux</i>	Km <sup>r</sup>	Jensch <i>et al.</i> , 2010
PN434	D39Δ <i>cps</i> Δ <i>spd</i> _0767	Km <sup>r</sup> , Spec <sup>r</sup>	This work
PN279	D39Δ <i>cps</i> Δ <i>spd</i> _0549	Km <sup>r</sup> , Erm <sup>r</sup>	This work
PN435	D39Δ <i>cps</i> Δ <i>spd</i> _0767Δ <i>spd</i> _0549	Km <sup>r</sup> , Erm <sup>r</sup> , Spec <sup>r</sup>	This work
PN438	D39/ <i>lux</i> Δ <i>spd</i> _0767	Km <sup>r</sup> , Spec <sup>r</sup>	This work
PN286	D39/ <i>lux</i> Δ <i>spd</i> _0549	Km <sup>r</sup> , Erm <sup>r</sup>	This work
PN439	D39/ <i>lux</i> Δ <i>spd</i> _0767Δ <i>spd</i> _0549	Km <sup>r</sup> , Erm <sup>r</sup> , Spec <sup>r</sup>	This work
PN220	D39Δ <i>cps</i> Δ <i>lgt</i>	Km <sup>r</sup> , Erm <sup>r</sup>	Pribyl <i>et al.</i> , 2014
<b><i>Escherichia coli</i></b>			
DH5α	Δ( <i>lac</i> )U169, <i>endA1</i> , <i>gyrA46</i> , <i>hsdR17</i> , Φ80Δ( <i>lacZ</i> )M15, <i>recA1</i> , <i>relA1</i> , <i>supE44</i> , <i>thi-1</i>	None	Bethesda Research Labs, Gaithersburg, U.S.
BL21(DE3)	<i>E. coli</i> B, <i>F</i> - <i>dcm ompT hsdS gal λ</i> (DE3), T7 polymerase gene under control of the <i>lacUV5</i> promoter	None	Novagen, Merck KGaA, Darmstadt, Germany
<b>Plasmids</b>			
pGEM <sup>®</sup> -T Easy	TA cloning vector for PCR products	Ap <sup>r</sup>	Madison, U.S.
pLITMUS28	multi-purpose cloning vector	Ap <sup>r</sup>	NEB, Frankfurt, Germany
pE89	pCR2.1Topo with erythromycin ( <i>ermB</i> ) cassette	Ap <sup>r</sup> , Km <sup>r</sup> , Erm <sup>r</sup>	Brehm <i>et al.</i> , 1987
pE96	pGEM-T with spectinomycin ( <i>aad9</i> ) cassette	Ap <sup>r</sup> , Spec <sup>r</sup>	LeBlanc <i>et al.</i> , 1991
p753	pLITMUS28 derivative with <i>sp</i> _0872 ( <i>dacA</i> ) 5' and 3' flanking region for mutagenesis	Ap <sup>r</sup>	This work
p574	pGEM-T derivative with <i>sp</i> _0629 ( <i>dacB</i> ) 5' and 3' flanking region for mutagenesis	Ap <sup>r</sup>	This work
P902	pLITMUS28 derivative with <i>sp</i> _0872 ( <i>dacA</i> ) interrupted by <i>aad9</i> resistance gene	Ap <sup>r</sup> , Spec <sup>r</sup>	This work
p647	pGEM-T derivative with <i>sp</i> _0629 ( <i>dacB</i> ) interrupted by <i>ermB</i> gene resistance cassette	Ap <sup>r</sup> , Erm <sup>r</sup>	This work
pET28	Protein expression vector	Km <sup>r</sup>	Novagen
pTP1	pET28 expression vector with a TEV protease cleavage site	Km <sup>r</sup> , Erm <sup>r</sup>	Saleh <i>et al.</i> , 2013
p923	pTP1 with D39 <i>spd</i> _0767 ( <i>dacA</i> ) for protein production and mice immunization	Km <sup>r</sup>	This work
p651	pTP1 with TIGR4 <i>sp</i> _0629 ( <i>dacB</i> ) for protein production and mice immunization ynd crystallization	Km <sup>r</sup>	This work
p947	pTP1 with TIGR4 <i>sp</i> _0629 ( <i>dacB</i> ) N-terminal truncated form for protein production and crystallization	Km <sup>r</sup>	This work
p949	pTP1 with TIGR4 <i>sp</i> _0629 ( <i>dacB</i> ) N-terminal truncated form for protein production and crystallization	Km <sup>r</sup>	This work

<sup>a</sup>Ap, ampicillin; Km, kanamycin; Erm, erythromycin; Spec, spectinomycin; r, resistant

Table S3. Primer list

Primer purpose	Primer	Sequence (5'-3')
<b>Insertion-deletion mutagenesis</b>		
Amplification of <i>sp_0872</i> + 5' and 3' flanking region	DacA_659 DacA_662	5'-CGGGATCCGATAAGAGCTGACAAGACG -3' 5'-CGCTCGAGGTCTAAATCCATCGCTAAAG -3'
Inverse PCR of <i>sp_0872</i> + 5' and 3' flanking region (pLITMUS28I)	DacA_661 DacA_660	5'-CGCCCCGGGTGGAATCAGTTTGTCCGC -3' 5'-CGCCCCGGGCTCAACAGCAATCGCATG -3'
Amplification of <i>sp_0629</i> + 5' and 3' flanking region	DacB_411 DacB_414	5'-CTACTAGAATTCTATGGGAGCCATGGGAGA-3' 5'-TTCCAAGCTGCAGTCGATCATACGCATTGCAG-3'
Inverse PCR of <i>sp_0629</i> + 5' and 3' flanking region (pGEM-T Easy)	DacB_413 DacB_412	5'-CTCACTGAAGCTTGGCTTTGAAGCGGAGA-3' 5'-ATCATCGAAGCTTCCATCAGCTTTGGCTGTC-3'
<b>Antibiotic resistance cassette amplification</b>		
erythromycin ( <i>ermB</i> )	ermB_105  ermB_106	5'-GATGATGATGATCCCGGTACCAAGCTTGAATTCACG GTTCTGTTCGTGCTG-3' 5'-AGTGAGTGAGTCCCGGGCTCGAGAAGCTTGA ATTCGTAGGCGCTAGGGACCTC-3'
spectinomycin ( <i>aad9</i> )	aad9_117 aad9_118	5'-AAAAAGCTTGAATTCGGATCCATCGATTTTCGTTCTGTAATAC-3' 5'-AAAAAGCTTGCTAGCAATTAGAATGAATATTTCCC-3'
<b>Recombinant protein production</b>		
<i>sp_0872</i> (TIGR4; <i>dacA</i> )	dacA_1068 dacA_1069	5'-GCGCGCTAGCCTAAGATTTTACCATTGCCGC -3' 5'-GGCCGAGCTCTATTTTCAATTTTCTGTCTGTACC -3'
<i>sp_0629</i> (TIGR4; <i>dacB</i> )	dacB_484 dacB_485	5'-GCGCGCTAGCCTAAGAAAAACAAAAATGAAG -3' 5'-GGCCAAGCTTAATCGACGTAGTCTCC -3'
<i>sp_0629</i> (TIGR4; <i>dacB</i> ) truncated I	dacB_1082 dacB_485	5'-GCGCGCTAGCCTGATTACTACAGCATTC-3' 5'-GGCCAAGCTTAATCGACGTAGTCTCC -3'
<i>sp_0629</i> (TIGR4; <i>dacB</i> ) truncated II	dacB_1140 dacB_485	5'-ATATGCTAGCGGAGCTGCCAGAAAGAAAG-3' 5'-GGCCAAGCTTAATCGACGTAGTCTCC -3'
<b>RT-PCR</b>		
<i>spd_0546</i> (D39)	SPD_0546_877 SPD_0546_878	5'-AACGCCAGTCTTTGCAATTT-3' 5'-TTGAGCTGTTGAGCCAAAGA-3'
<i>spd_0547</i> (D39)	SPD_0547_868	5'-AGTCTCGCCATTGGATGAC-3'
<i>spd_0548</i> (D39)	SPD_0548_870 SPD_0548_871	5'-TGGAACAGGCTATCTTGTGG-3' 5'-ACGACCCTTGAGACCATGAC-3'
<i>spd_0549</i> (D39)	SPD_0549_872	5'-AGGAGCTGCCCAGAAGAAAG-3'
<b>Northern blotting</b>		
<i>spd_0547</i> (D39)	SPD_0547_943 SPD_0547_944	5'-TGCGTCGCCAACAGGCTTTA-3' 5'-CTAATACGACTCACTATAGGGAGATTTGGTGGATGAGGATGGTT-3'
<i>spd_0549</i> (D39)	SPD_0549_941 SPD_0549_942	5'-TCTTGACAGCTTTGCTAGCCTTGAG-3' 5'-CTAATACGACTCACTATAGGGAGATGGGCTGCTTTTCTTCTGT-3'

<sup>a</sup>Restriction sites are underlined

---

## Legends to supplementary Figures

**Fig. S1.** Growth behavior of *S. pneumoniae* parental strains D39 and D39 $\Delta$ *cps* and of their isogenic  $\Delta$ *dacA*,  $\Delta$ *dacB*, and  $\Delta$ *dacA* $\Delta$ *dacB* mutants.

Pneumococcal wild-type D39 and D39 $\Delta$ *cps* and the isogenic mutants  $\Delta$ *dacA*,  $\Delta$ *dacB*, and  $\Delta$ *dacA* $\Delta$ *dacB* were cultured at 37°C and 5% CO<sub>2</sub> either in complex THY medium or chemically-defined medium (CDM): The appropriate antibiotics (spectinomycin; 50 µg/ml) or (erythromycin; 2.5 µg/ml) were added for the mutants. Growth of the  $\Delta$ *dacA*,  $\Delta$ *dacB*, and  $\Delta$ *dacA* $\Delta$ *dacB* mutants was moderately delay compared to the encapsulated D39 parental strain when bacterial growth was monitored in THY medium. This effect was more pronounced in CDM. The nonencapsulated *dac*-mutants showed already a higher tendency to lyse in THY. In CDM the nonencapsulated *dac*-mutants lyse immediately after reaching the stationary phase and did not reach the bacterial density as the D39 $\Delta$ *cps* parental strain.

**Fig. S2.** Structures and calculated masses for the identified monomeric LytA-released pneumococcal PGN-peptides.

**Fig. S3.** Structures and calculated masses for the identified dimeric LytA-released pneumococcal PGN-peptides.

**Fig. S4.** ESI-FT-ICR-MS spectra (charge deconvoluted) of PGN-peptides isolated from strain *S.p.* D39 $\Delta$ *cps*.

Spectrum of P-30 pool 1 (A) and spectrum of P-30 pool 2 (B). Mass accuracy is given in ppm, peak assignment is based on HPLC profiles (Fig. S9 and S10) and MS/MS experiments. The majority of identified peptides clearly indicates the WT PGN-type: **1** (as Na<sup>+</sup>-adduct ion), **4**, **4'**, **9**, **12A/12B**, **12A/12B'**, **17**, and **17'/17'<sup>§</sup>**. Peptides **2** and **3\*** are derived from not fully processed PGN-peptide strands.

---

**Fig. S5.** ESI-FT-ICR-MS spectra (charge deconvoluted) of PGN-peptides isolated from strain *S.p.* D39 $\Delta$ *cps* $\Delta$ *dacA*.

Spectrum of P-30 pool 1 (A) and spectrum of P-30 pool 2 (B). Mass accuracy is given in ppm, peak assignment is based on HPLC profiles (Fig. S9 and S10) and MS/MS experiments. Most of the identified PGN-peptides clearly display the  $\Delta$ *dacA*-PGN type, because a pentapeptid was detected as stem peptide: **3**, **3\***, **8**, **11**, **11\***, **16**, **16\***, **16'**, and **16'\***. In addition some peptides from WT PGN have been observed: **9**, **12A/12B**, and **12A'/12B'**.

**Fig. S6.** ESI-FT-ICR-MS spectra (charge deconvoluted) of PGN-peptides isolated from strain *S.p.* D39 $\Delta$ *cps* $\Delta$ *dacA* $\Delta$ *dacB*.

Spectrum of P-30 pool 1 (A) and spectrum of P-30 pool 2 (B). Mass accuracy is given in ppm, peak assignment is based on HPLC profiles (Fig. S9 and S10) and MS/MS experiments. The  $\Delta$ *dacA* $\Delta$ *dacB*-mutant shows mainly a  $\Delta$ *dacA*-PGN type because the same molecules with a pentapeptide as stem peptide were identified: **3**, **3\***, **8**, **11**, **11\***, **16**, **16\***, and **16'\*** (compare Fig. S5) as well as **7**, **15**, and **15'** in addition. Peptides **2** and **10** belong to the L,D-CPase impaired ( $\Delta$ *dacB*) PGN-type. (# peaks could not be assigned to a PGN-peptide; ## peak can originate from TetraTetra(SA)<sub>2</sub> or Penta(SA)Tetra(S))

**Fig. S7.** ESI-FT-ICR-MS spectrum (charge deconvoluted) of PGN-peptides isolated from strain *S.p.* D39 $\Delta$ *cps* $\Delta$ *dacB* P-30 pool 2.

The spectrum of the respective P-30 pool 1 is depicted in Fig. 4C. Mass accuracy is given in ppm, peak assignment is based on HPLC profiles (Fig. S9 and S10) and MS/MS experiments. All identified peptides clearly indicate an L,D-CPase impaired PGN-type since they have a tetrapeptide as their stem: **2**, **5/5'**, **10**, **13/13'**, **14/14'**, and **18/18'** (and respective Na<sup>+</sup>-adduct ions and/or ions for 2M).

---

**Fig. S8.** ESI-FT-ICR-MS spectrum (charge deconvoluted) of PGN-peptides isolated from strain *S.p.* D39 $\Delta$ *cps* $\Delta$ *lgt* P-30 pool 2.

The spectrum of the respective P-30 pool 1 is depicted in Fig. 4D. Mass accuracy is given in ppm, peak assignment is based on HPLC profiles (Fig. S9 and S10) and MS/MS experiments (# signals derive from cluster of **1** (M) with a second molecule). The analysis of the peptides isolated from the  $\Delta$ *lgt*-mutant clearly reveals that this mutant strain displays a WT PGN-type in contrast to the some extend expected  $\Delta$ *dacB*-PGN-type. The peptides indicative for the WT PGN-type are: **1** (as monomer, dimer, trimer, and cluster ions), **4**, **4'**, **9**, **12A/12B**, **12A/12B'**, **17**, and **17'/17'**<sup>§</sup>.

**Fig. S9.** RP-HPLC profiles of the PGN-peptide GPC pools 2, which mostly contain the monomeric peptides. A similarity between the profiles of D39 $\Delta$ *cps* and D39 $\Delta$ *cps* $\Delta$ *lgt* is observed as well as between D39 $\Delta$ *cps* $\Delta$ *dacA* and D39 $\Delta$ *cps* $\Delta$ *dacA* $\Delta$ *dacB*. The profile for D39 $\Delta$ *cps* $\Delta$ *dacB* shows no clear similarity to either profiles described above. In the profile of D39 $\Delta$ *cps* $\Delta$ *dacA* some WT PGN-peptides were identified. Chromatograms were recorded at 210 nm, peak assignment is based on MS(/MS) analysis. For the respective peptide structures see figures S2 and S3.

**Fig. S10.** RP-HPLC profiles of the PGN-peptide GPC pools 1, which mostly contain peptide dimers. A similarity between the profiles for D39 $\Delta$ *cps* and D39 $\Delta$ *cps* $\Delta$ *lgt* is observed as well as between D39 $\Delta$ *cps* $\Delta$ *dacA* and D39 $\Delta$ *cps* $\Delta$ *dacA* $\Delta$ *dacB*. The profile of D39 $\Delta$ *cps* $\Delta$ *dacB* exhibits a higher similarity to the latter ones, because the major peptides have almost the identical retention times (**10** compared to **11**, **14/14'** compared to **16/16'**). But the peptides are clearly different entities, which can easily be shown by mass spectrometry. Chromatograms were recorded at 210 nm, peak assignment is based on MS(/MS) analysis. For peptide structures see figures S2 and S3.

---

**Fig. S11.** ESI-MS/MS analysis of **3\*** (A) compared to **3** (B), clearly showing that a glycine is present instead of an alanine at position 5 of the stem peptide.

**Fig. S12.** Purification and immunoblot analysis of recombinant DacA and DacB proteins.

DacA (**A**) and DacB (**B**) protein expression was induced with IPTG and total protein lysates were subjected to SDS-PAGE followed by coomassie brilliant blue (CCB) staining. The His<sub>6</sub>-tagged proteins (His<sub>6</sub>-DacA and His<sub>6</sub>-DacB) were purified by affinity chromatography and the His<sub>6</sub>-tag was removed using TEV protease as described (Saleh *et al.*, 2013). The proteins were detected by immunoblot analysis using anti-Dac antibodies. NI, non-induced; IN, induced; rP, recombinant protein.

**Fig. S13:** Sequence and secondary structure elements of *S. pneumoniae* DacB.

Sequence of *S. pneumoniae* DacB showing the secondary structure elements, the  $\beta 5$ - $\alpha 4$  loop is colored in green and residues involved in coordination of catalytic Zn<sup>2+</sup> are inserted in red boxes. Secondary structural elements are represented for chain A (chain B presents  $\alpha 4$  helix with four residues more).

**Fig. S14:** Structural comparison of DacB (chain A) with its closest homologues.

The crystal structures of BaLdcB (cyan, PDB code 4MPH), VanXYc (grey, PDB code 4MUS) and VanXYg (white, PDB code 4F78) are superimposed with DacB (orange, PDB code 4D0Y). The  $\beta 5$ - $\alpha 4$  loop is different between D,D-carboxypeptidases (VanXYc and VanXYg) and L,D-carboxypeptidases (DacB and BaLdcB). In BaLdcB, VanXYc and VanXYg enzymes there is an extra-loop before  $\alpha 1$  that is not present in DacB.

**Fig. S15:** Structural comparison between *S. pneumoniae* DacB and DacA.

(**A**) Overall structure and catalytic site composition in DacB. Zn<sup>2+</sup> atom is shown in red. Relevant residues conforming the DacB active site are labeled and represented in sticks. Phosphate molecule attached to the active site of DacB is represented in sticks.

---

**(B)** Overall structure of DacA. Relevant residues conforming the DacA active site are labeled and represented in sticks. Sulphate molecule attached to the active site of DacA is represented in sticks. The critical omega-like loop is labeled.

**Fig. S16:** Electrostatic potential on the molecular surfaces of DacB and DacA.

Electrostatic potential on the molecular surface is colored in blue for basic regions and in red for acidic regions. The substrate-binding sites in DacB and DacA are indicated by dashed boxes. DacB presents a more acidic character in agreement with recognition of the basic L-Lys residue while DacA presents a strong basic character that would recognize the carboxy terminal moiety and, at the same time, would impede the interaction of the L-Lys residue from the peptide stem.

**Fig. S17:** Structural comparison of loop  $\beta 5\text{-}\alpha 4$  in *S. pneumoniae* DacB and *E. gallinarum* VanXYc.

Superposition of the closed conformation of DacB (chain A, colored in yellow) with the open conformation of DacB (chain B, colored in blue) and VanXYc in complex with D-Ala-D-Ala (PDB code 4OAK, in grey). The D-Ala-D-Ala substrate is represented in pink sticks. While VanXYc shows a long loop covering the catalytic cavity, DacB presents two different conformations for the loop providing two different conformations for the active. Catalytic cations are represented as spheres.

**Fig. S18:** Different structural rearrangement for  $\beta 5\text{-}\alpha 4$  loop between the open (blue) and closed (yellow) DacB conformations.

Superposition of DacB in the open conformation (chain B, blue) and in the closed one (chain A, yellow). Residues involved in  $\text{Zn}^{2+}$  coordination, the residues of the  $\beta 5\text{-}\alpha 4$  loop and the Trp206 are represented in sticks. Important structural changes are observed between the open and closed conformation, the largest change corresponding to Leu168 (dashed line).

---

The loop rearrangement also affects the side chain conformation of Trp206, which rotates about 90° to approach the active site.



Figure S1

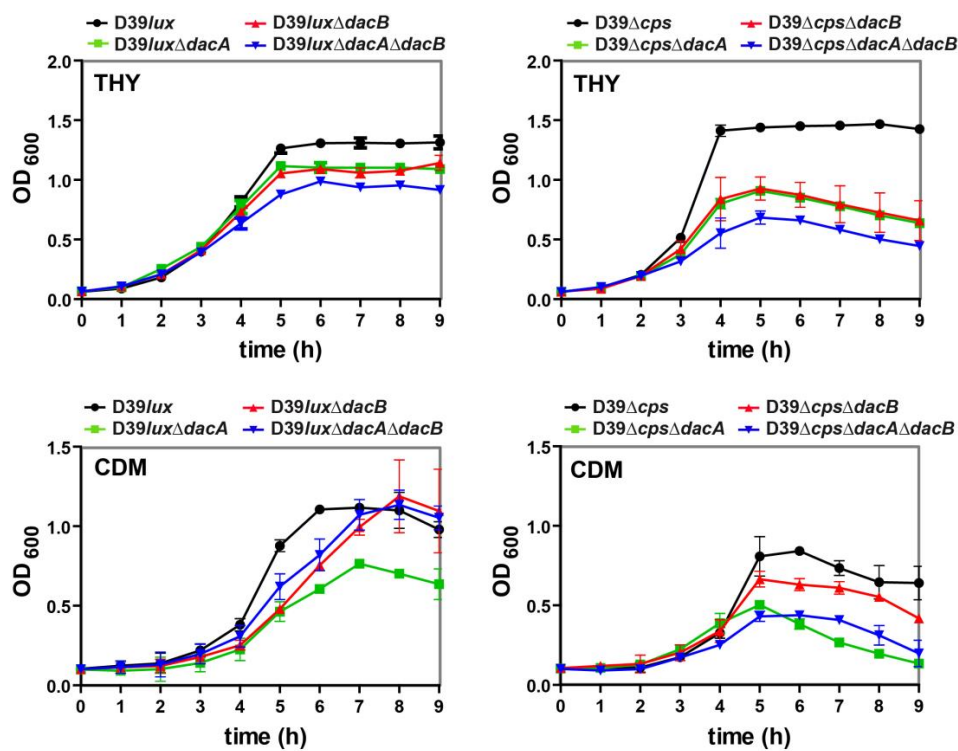
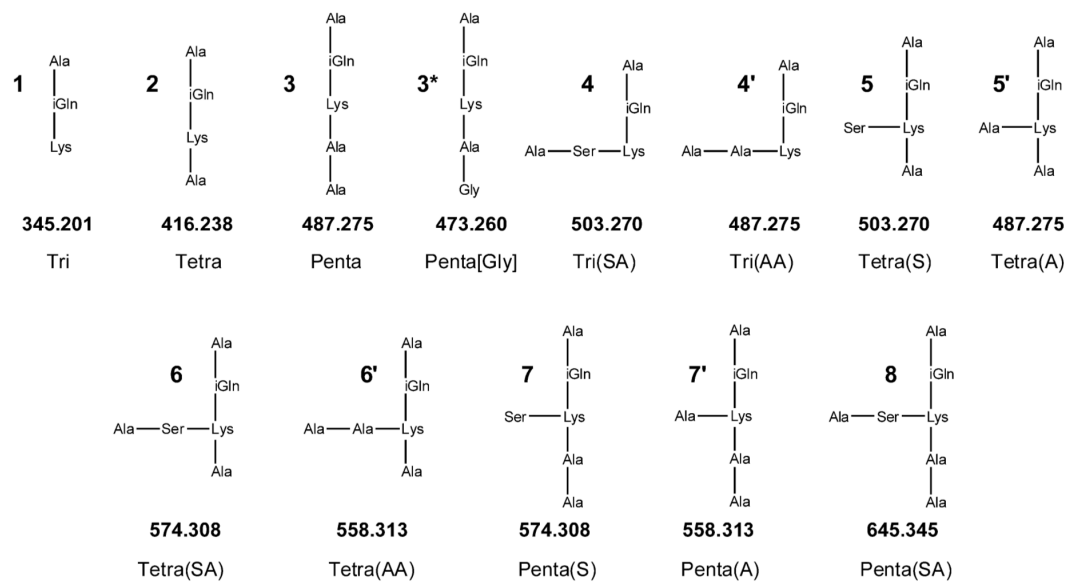


Figure S2



**Figure S3**

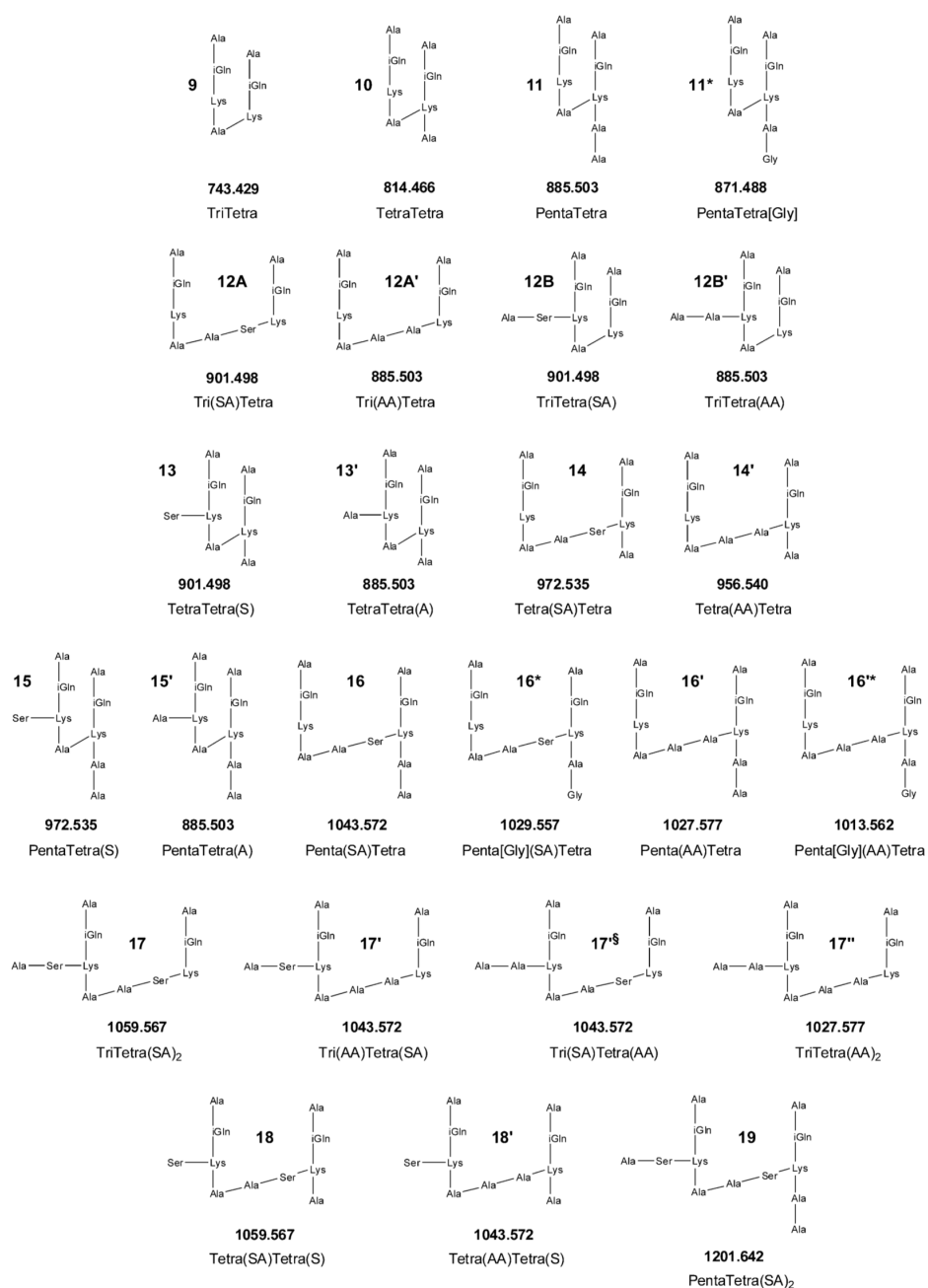


Figure S4

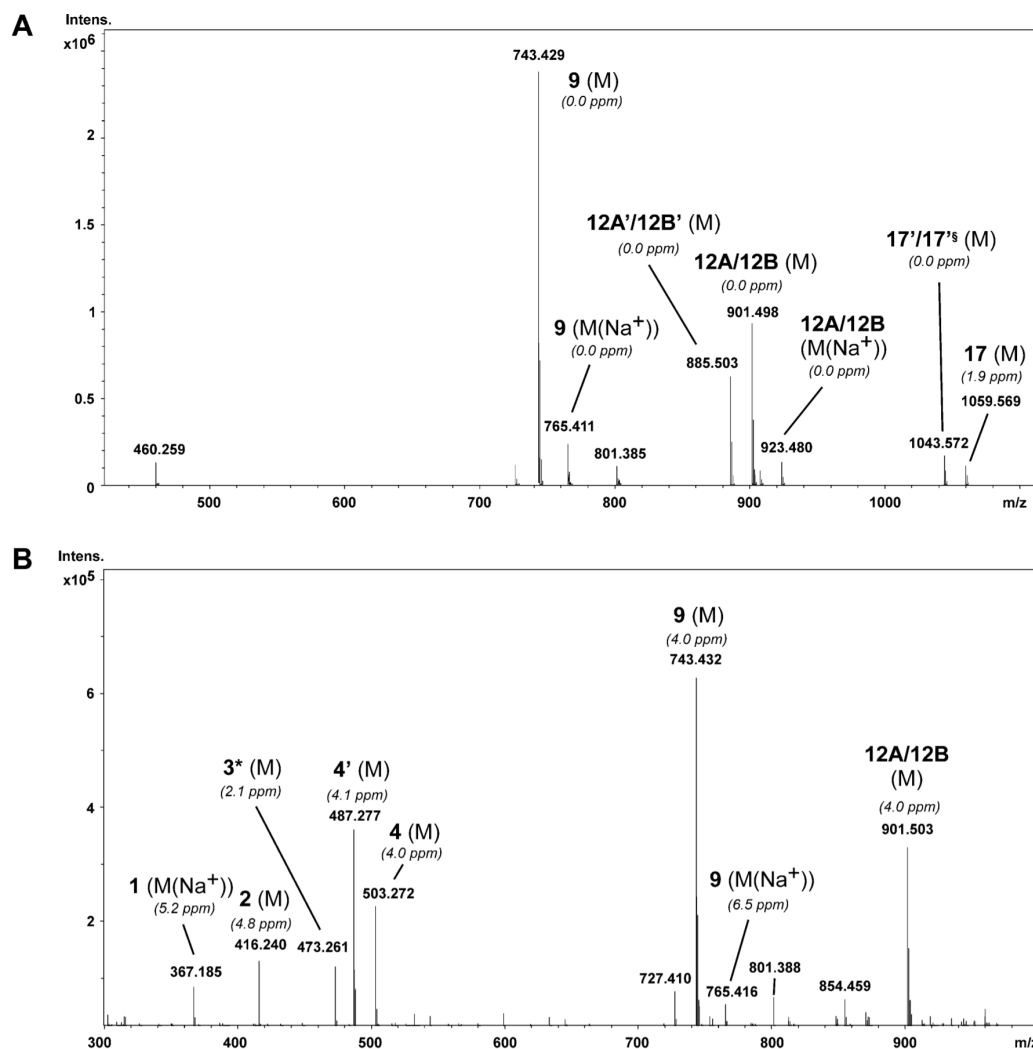


Figure S5

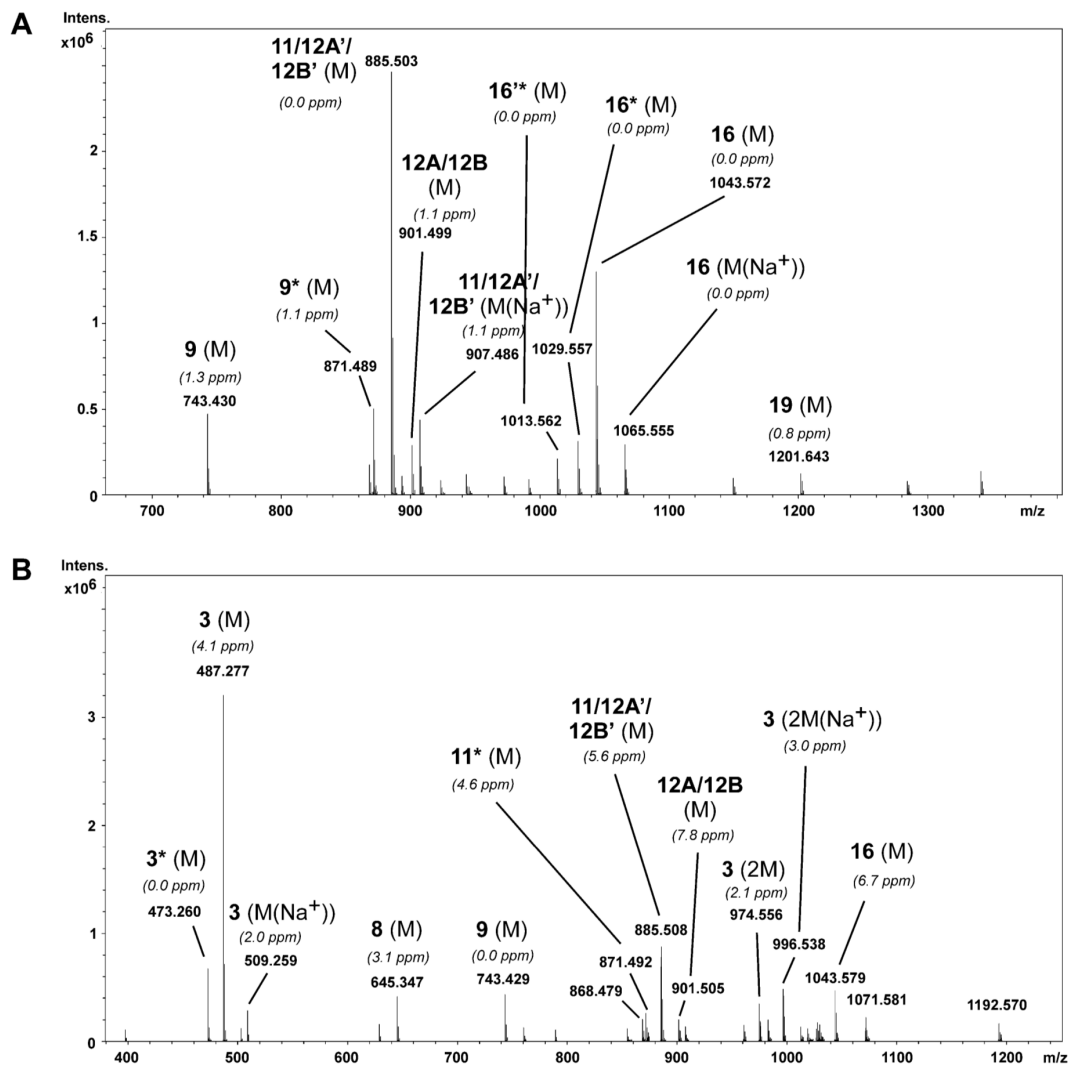




Figure S7

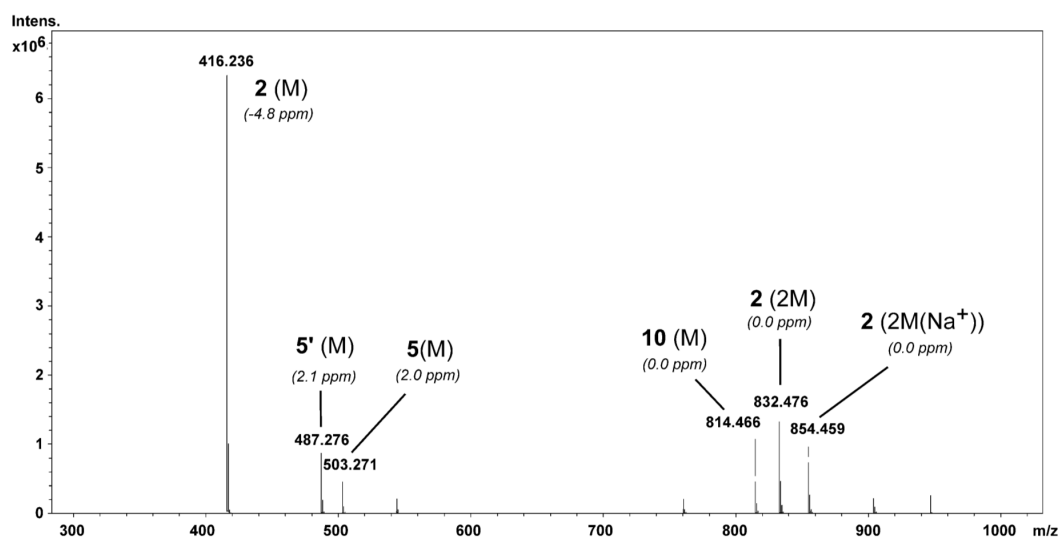


Figure S8

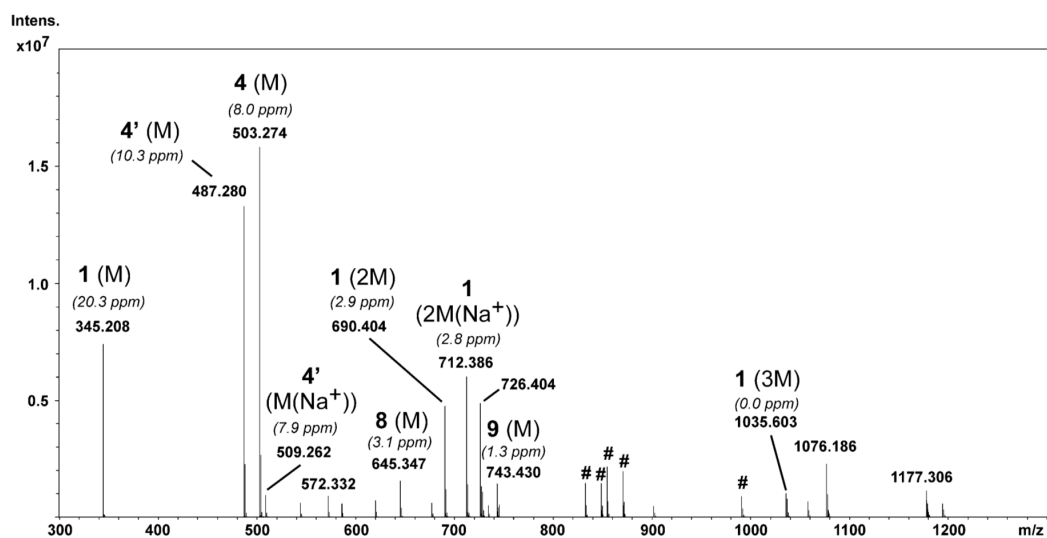


Figure S9

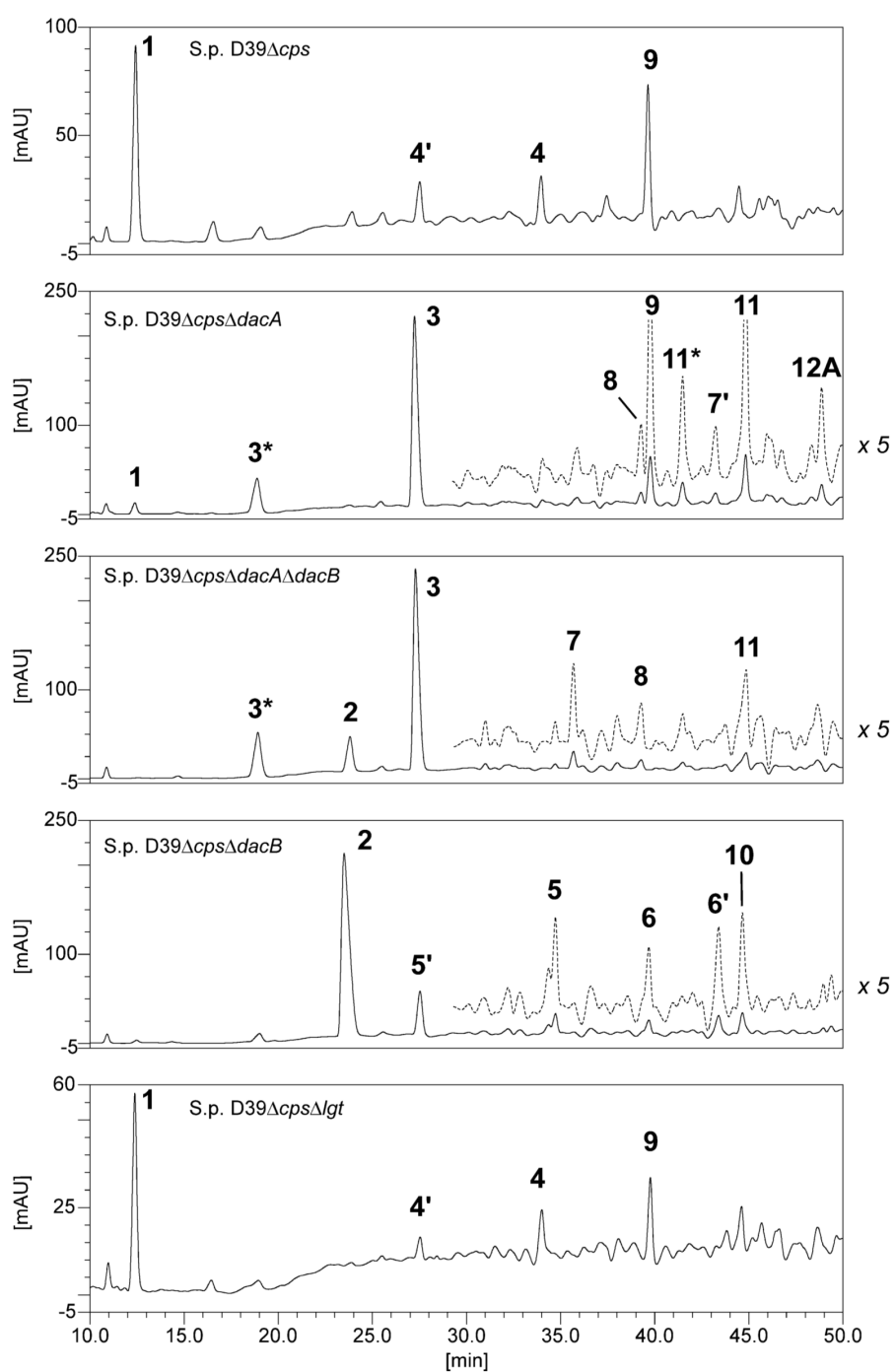




Figure S10

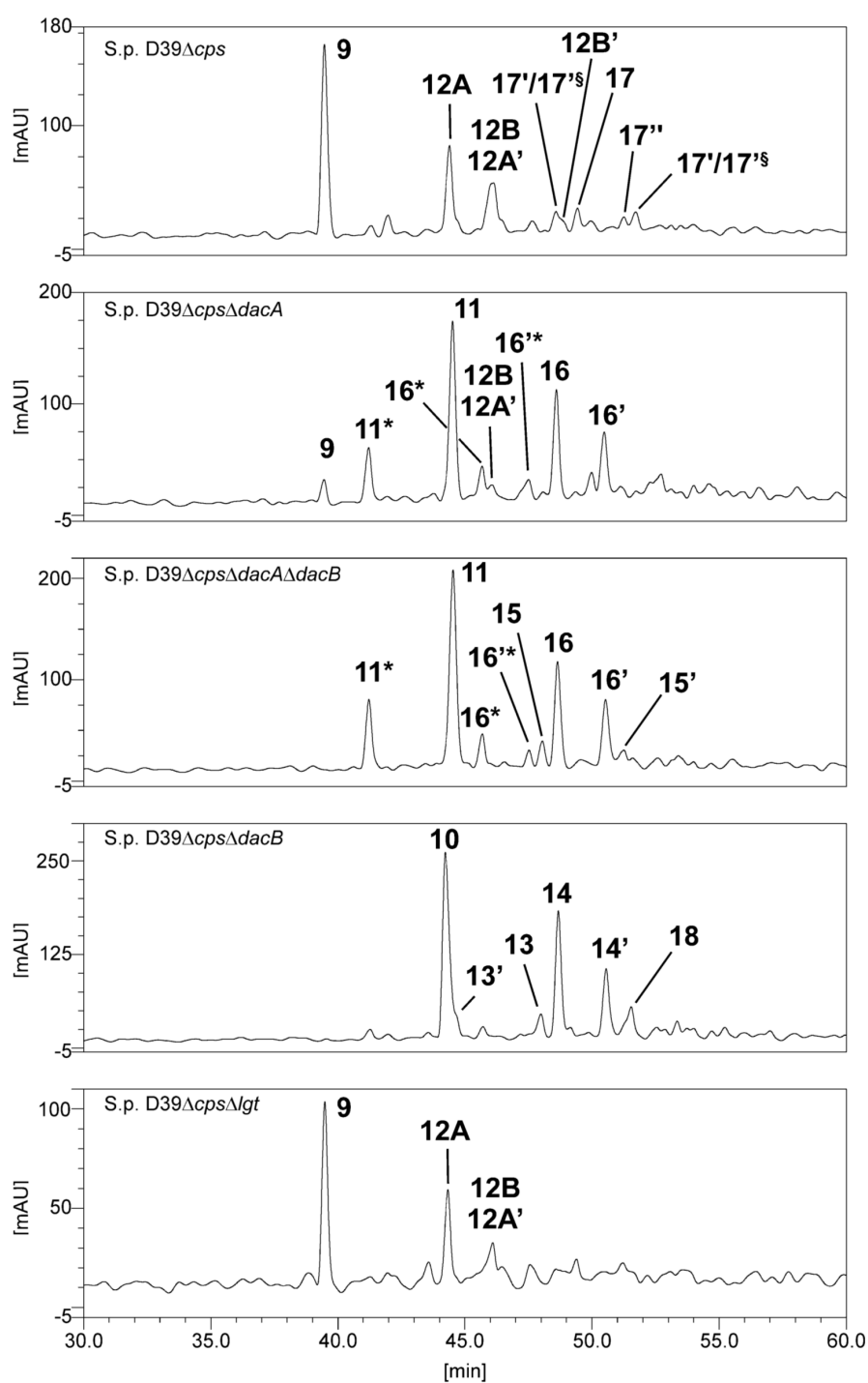


Figure S11

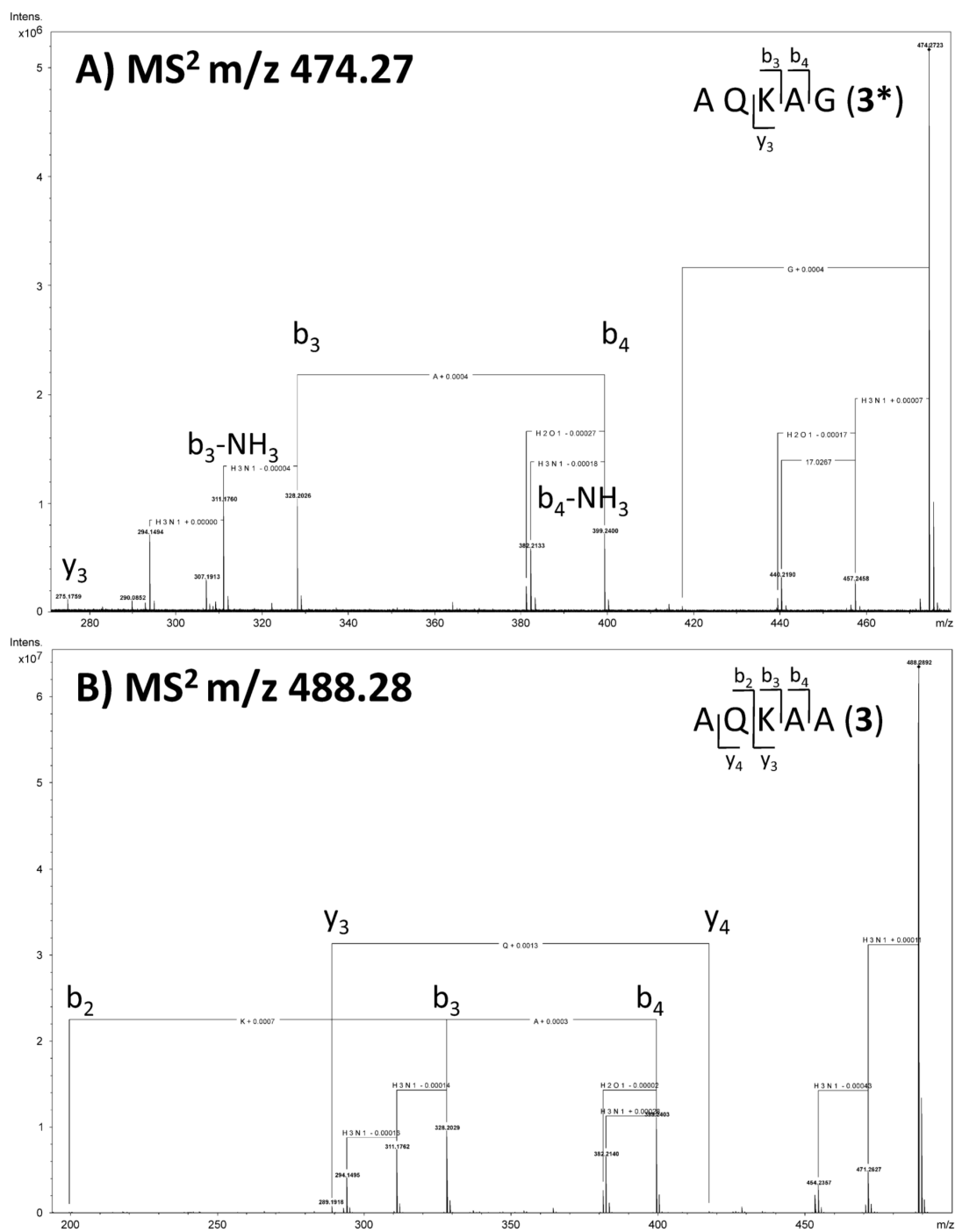


Figure S12

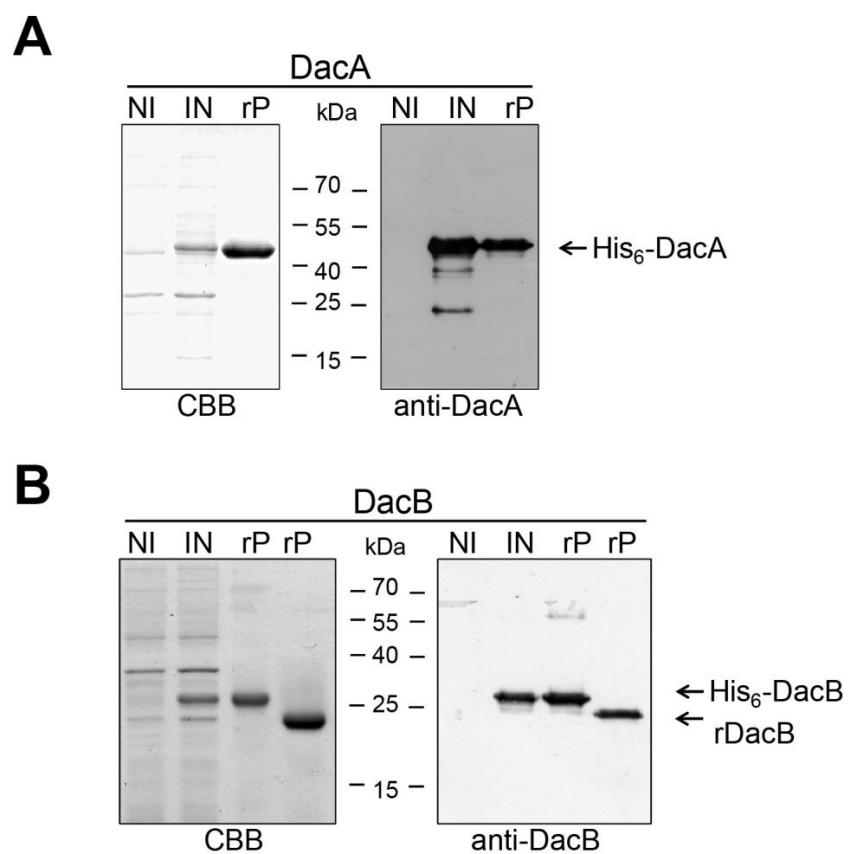
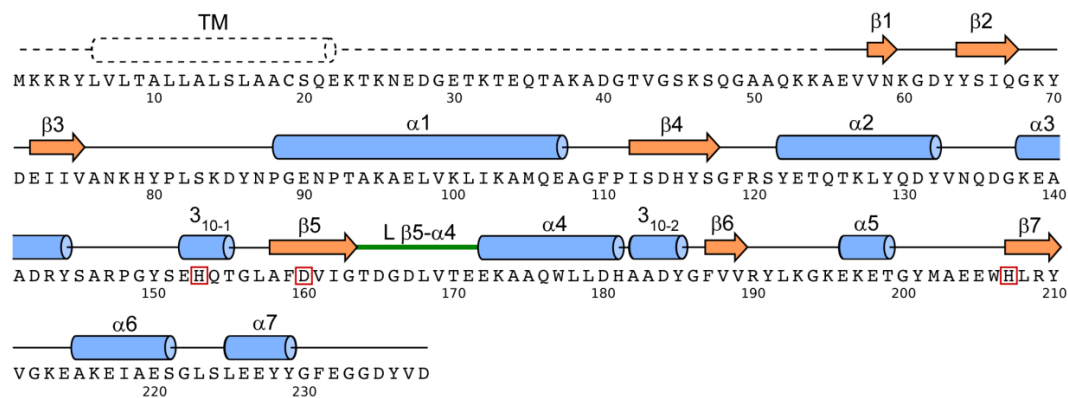


Figure S13



---

Figure S14

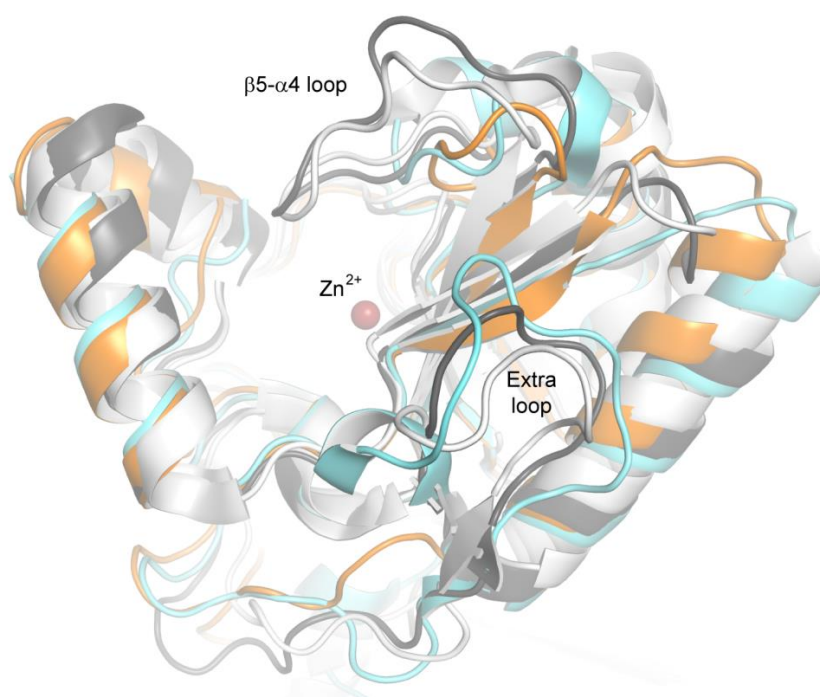
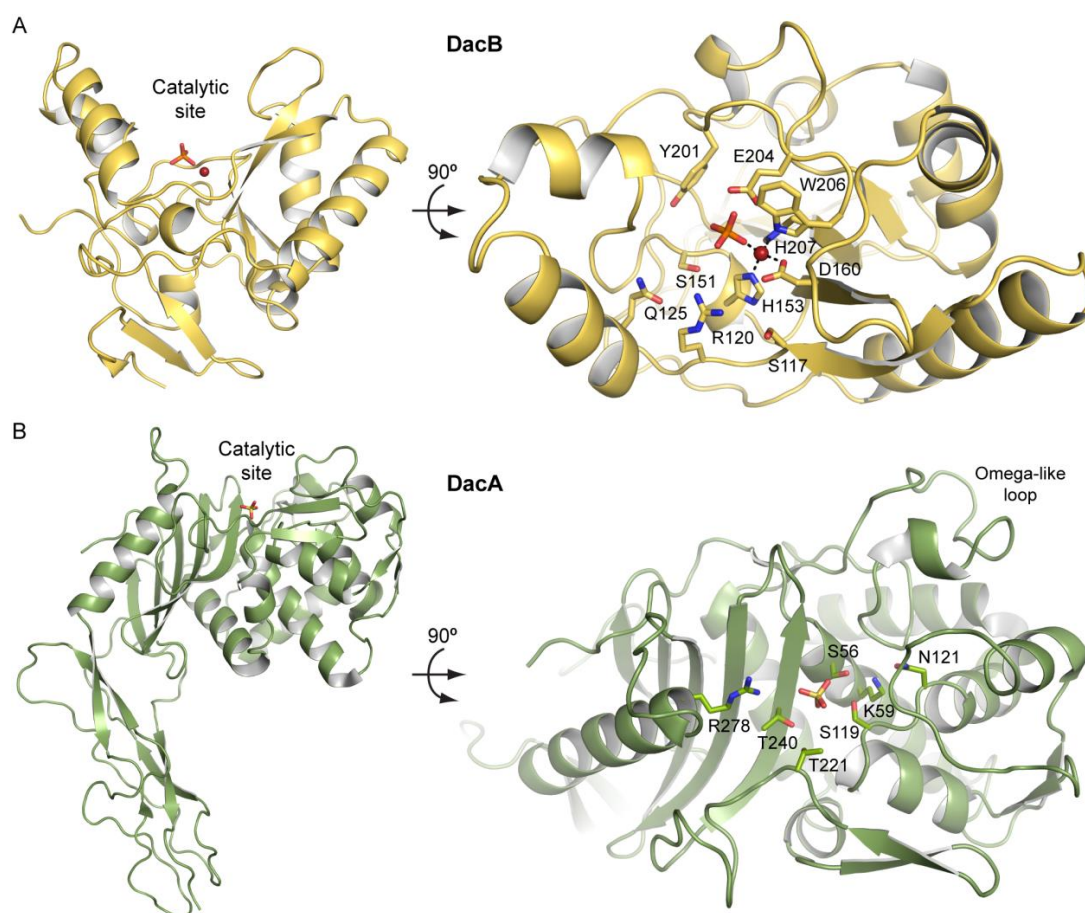


Figure S15



---

Figure S16

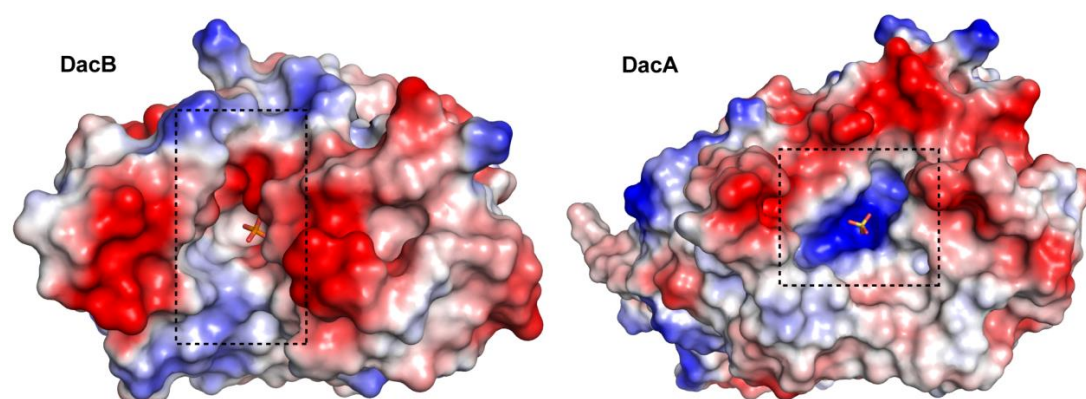


Figure S17

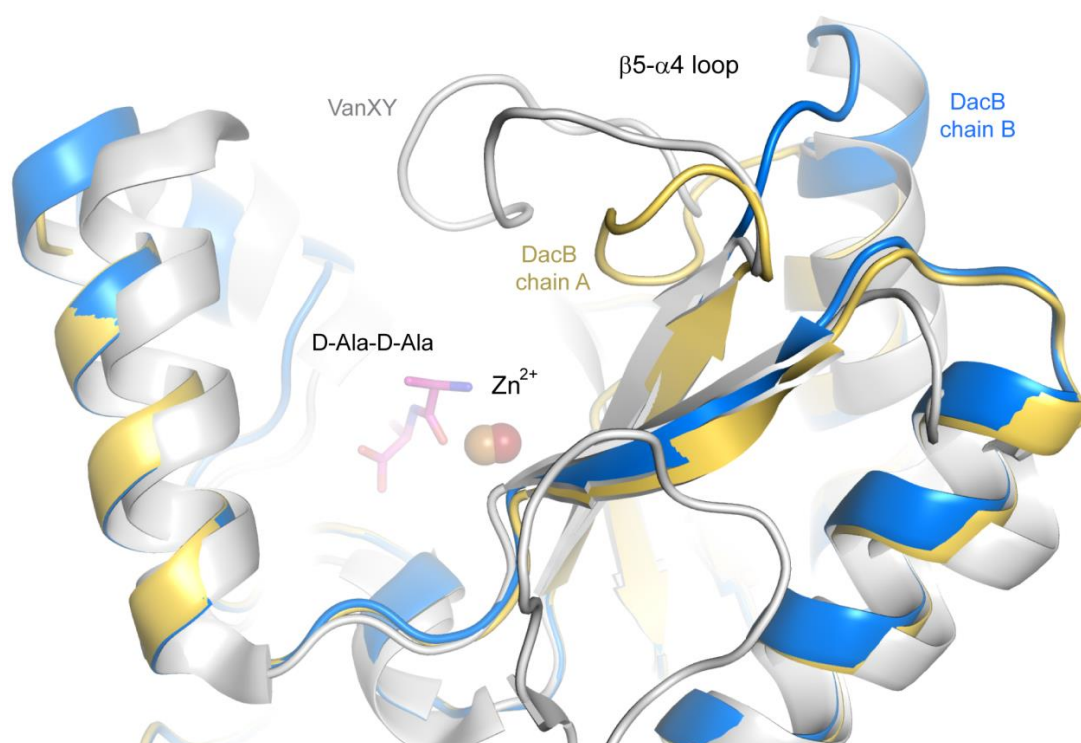




Figure S18

